



TUGAS AKHIR - SS141501

**PERAMALAN CURAH HUJAN SEBAGAI PENDUKUNG
KALENDER TANAM PADI DI KABUPATEN
BOJONEGORO MENGGUNAKAN METODE ARIMA,
SUPPORT VECTOR REGRESSION DAN *GENETIC
ALGORITHM-SUPPORT VECTOR REGRESSION***

**KIKI WULAN SUCI
NRP 1313 100 086**

**Dosen Pembimbing
Irhamah, M.Si, Ph.D**

**PROGRAM STUDI S1
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



TUGAS AKHIR - SS 141501

**PERAMALAN CURAH HUJAN SEBAGAI PENDUKUNG
KALENDER TANAM PADI DI KABUPATEN BOJONEGORO
MENGUNAKAN METODE ARIMA, *SUPPORT VECTOR
REGRESSION* DAN *GENETIC ALGORITHM-SUPPORT
VECTOR REGRESSION***

**KIKI WULAN SUCI
NRP 1313 100 086**

**Dosen Pembimbing:
Irhamah, M.Si, Ph.D**

**PROGRAM STUDI S1
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**



FINAL PROJECT - SS 141501

**RAINFALL FORECASTING TO SUPPORT RICE PLANTING
CALENDER IN BOJONEGORO DISTRICT USING ARIMA,
*SUPPORT VECTOR REGRESSION AND GENETIC
ALGORITHM-SUPPORT VECTOR REGRESSION***

KIKI WULAN SUCI
NRP 1313 100 086

Supervisor:
Irhamah, M.Si, Ph.D

**UNDERGRADUATE PROGRAMME
DEPARTEMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2017**

LEMBAR PENGESAHAN

**PERAMALAN CURAH HUJAN SEBAGAI PENDUKUNG
KALENDER TANAM PADI DI KABUPATEN
BOJONEGORO MENGGUNAKAN METODE ARIMA,
SUPPORT VECTOR REGRESSION DAN GENETIC
ALGORITHM-SUPPORT VECTOR REGRESSION**

TUGAS AKHIR

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains
pada

Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember

Oleh :

KIKI WULAN SUCI
NRP 1313 100 086

Disetujui oleh Pembimbing Tugas Akhir

Irhamah, M.Si, Ph.D
NIP. 19780406 200112 2 002

(*Irhamah*)

Mengetahui
Ketua Jurusan Statistika FMIPA-ITS



Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JANUARI 2017

**PERAMALAN CURAH HUJAN SEBAGAI PENDUKUNG
KALENDER TANAM PADI DI KABUPATEN
BOJONEGORO MENGGUNAKAN METODE ARIMA,
SUPPORT VECTOR REGRESSION DAN GENETIC
ALGORITHM-SUPPORT VECTOR REGRESSION**

**Nama Mahasiswa : Kiki Wulan Suci
NRP : 1313 100 086
Jurusan : Statistika
Pembimbing : Irhamah, M.Si, Ph.D**

ABSTRAK

Kabupaten Bojonegoro dikenal sebagai lumbung padi di Jawa Timur. Sebanyak 33,31% lahan Bojonegoro digunakan sebagai lahan sawah. Pada beberapa tahun terakhir produksi padi Kabupaten Bojonegoro mengalami fluktuasi yang salah satu penyebabnya adalah iklim ekstrem. Iklim ekstrem dapat mengakibatkan efek musim kemarau yang panjang serta adanya hujan ekstrem yang mengakibatkan petani mengalami kerugian. Oleh karena itu, keberhasilan produksi padi sangat bergantung pada informasi mengenai data curah hujan yang tersusun dalam kalender tanam. Dalam penelitian ini dilakukan peramalan curah hujan dasarian di Pos Cawak dan Kedungadem menggunakan metode ARIMA, Support Vector Regression (SVR) dan Genetic Algorithm-SVR (GA-SVR). Berdasarkan RMSE dan SMAPE metode GA-SVR menghasilkan peramalan yang lebih akurat. Berdasarkan forecast 6 bulan selanjutnya akan dibuat kalender tanam padi. Hasil kalender tanam padi pada bulan Juli 2016- Desember 2016 menunjukkan kebutuhan air untuk penanaman padi sawah tidak dapat terpenuhi. Petani dapat mengganti padi dengan menanam palawija. Jika tetap menanam padi, maka petani dan pemerintah harus memastikan tersedianya cadangan air.

Kata Kunci : ARIMA, Curah Hujan, Genetic Algorithm, Support Vector Regression

(Halaman ini sengaja dikosongkan)

**RAINFALL FORECASTING TO SUPPORT RICE
PLANTING CALENDAR IN BOJONEGORO DISTRICT
USING ARIMA, SUPPORT VECTOR REGRESSION AND
GENETIC-ALGORITHM SUPPORT VECTOR REGRESSION**

Name	: Kiki Wulan Suci
NRP	: 1313 100 086
Department	: Statistika
Supervisor	: Irhamah, M.Si, Ph.D

ABSTRACT

Bojonegoro District known as a granary in East Java. A total of 33.31% of the land in Bojonegoro is used as a wetland. In recent years the production of rice in Bojonegoro fluctuated, which one of them caused by extreme climate. Extreme climate can caused continuously dry season and extreme rainfall can caused farmers suffer losses. Therefore, the success of rice production is highly dependent on information of the rainfall data which arranged in a planting calendar. In this study, rainfall forecasting is done per 10 days in Cawak and Kedungadem Station using ARIMA method, Support Vector Regression (SVR) and Genetic Algorithm-SVR (GA-SVR). Based on RMSE and SMAPE values, GA-SVR method gave better forecast accuracy. Rice planting calendar will be made based on the forecasting result of the next 6 months. The result of rice planting calendar on July 2016 – December 2016 indicates that the water requirement for rice cultivation cannot be fulfilled. Farmers may substitute it with palawija which require less water, but if they keep doing the rice planting, the farmers and government must ensure the availability of water reserves.

Keyword : ARIMA, Genetic Algorithm, Rainfall, Support Vector Regression

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur kehadapan Tuhan Yang Maha Esa karena atas anugrah-Nya sehingga penulis dapat menyelesaikan dengan baik tugas akhir dengan judul **Peramalan Curah Hujan Sebagai Pendukung Kalender Tanam Padi di Kabupaten Bojonegoro Menggunakan Metode ARIMA, *Support Vector Regression* dan *Genetic Algorithm-Support Vector Regression***. Proses penulisan tugas akhir ini dapat diselesaikan dengan baik dengan bantuan banyak pihak, oleh karena itu penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu yaitu :

1. Bapak Dr. Suhartono selaku Ketua Jurusan Statistika dan Bapak Dr. Sutikno, M.Si selaku Ketua Prodi S1 Statistika.
2. Ibu Irhamah, M.Si, Ph.D selaku dosen pembimbing yang telah dengan sabar memberikan arahan dan bimbingan dalam penulisan tugas akhir ini.
3. Bapak Dr.rer.pol Heri Kuswanto, M.Si dan Ibu Dra. Wiwiek Setya Winahyu, MS selaku dosen penguji yang telah memberikan kritik dan saran dalam penyempurnaan tugas akhir ini.
4. Bapak Dr. Sutikno, M.Si selaku dosen wali atas nasehat dan semangat yang diberikan.
5. Seluruh Bapak dan Ibu dosen Jurusan Statistika yang telah memberikan banyak ilmu pengetahuan dibidang statistika dan penerapannya.
6. Pihak BMKG Karangploso Malang yang telah memberikan izin untuk memperoleh data curah hujan dasaharian di Pos Cawak dan Pos Kedungadem Bojonegoro.
7. Kedua orang tua penulis, Khayin Abdad dan Siti Aisyah atas doa, semangat dan dukungan hingga tugas akhir ini selesai.
8. Keluarga besar Khayin Abdad yang telah banyak memberikan bantuan dan motivasi dalam proses penyelesain tugas akhir ini.
9. Keluarga Bani Ali Astikah yang telah banyak memberikan bantuan dan motivasi dalam proses penyelesain tugas akhir ini.

10. Krisna, Ageng, Adam, Elchemy, Mas Risma, Mas Rizky dan Mas Akhmad yang telah membantu, memberikan dukungan, saran, motivasi selama proses pengerjaan Tugas Akhir ini.
11. Teman-teman angkatan 2013, LEGENDARY jurusan statistika yang telah memberikan dukungan dan bantuan.
12. Teman-teman Maichi yang telah memberikan semangat dalam proses penyelesaian tugas akhir ini.
13. Teman-teman anggota Divisi PSt HIMASTA-ITS yang telah membagi ilmu dan memberi kesempatan untuk berkarya dalam bidang nonakademis.

Penulis juga mengucapkan terima kasih kepada semua pihak yang tidak dapat disebutkan satu persatu. Semoga bantuan yang telah diberikan dalam penyusunan tugas akhir ini mendapat balasan dari Tuhan Yang Maha Esa.

Surabaya, Januari 2017

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xvii
DAFTAR GAMBAR	xxiii
DAFTAR LAMPIRAN	xxvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Tujuan	5
1.4 Manfaat	6
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA.....	7
2.1 <i>Autoregressive Integrated Moving Average</i> (ARIMA).....	7
2.1.1 Stasioneritas Model ARIMA	7
2.1.2 Identifikasi Model ARIMA	9
2.1.3 Estimasi Parameter Model ARIMA	10
2.1.4 Pengujian Signifikansi Parameter	14

2.1.5	<i>Diagnostic Checking</i>	15
2.2	Uji Nonlinieritas Terasvirta	15
2.3	<i>Support Vector Regression</i> (SVR)	16
2.4	<i>Genetic Algorithm</i>	19
2.5	Kriteria Pemilihan Model Terbaik	21
2.6	Curah Hujan	22
2.7	Penentuan Pola Tanam Padi dan Kalender Tanam Padi	22
BAB III METODOLOGI PENELITIAN		25
3.1	Sumber Data	25
3.2	Variabel Penelitian	25
3.3	Langkah Analisis	26
3.4	Diagram Alir	29
BAB IV ANALISIS DAN PEMBAHASAN		33
4.1	Karakteristik Curah Hujan Kabupaten Bojonegoro	33
4.1.1	Deskripsi Curah Hujan Pos Cawak.....	34
4.1.2	Deskripsi Curah Hujan Pos Kedungadem	35
4.2	Peramalan Curah Hujan Pos Cawak.....	37
4.2.1	Peramalan Curah Hujan Menggunakan ARIMA	37
4.2.2	Peramalan Curah Hujan Menggunakan SVR .	54
4.2.3	Peramalan Curah Hujan Menggunakan GA SVR	59
4.2.4	Perbandingan Model Terbaik Pos Cawak.....	65

4.2.5	Kalender Tanam Padi Pos Cawak Periode Juli-Desember Tahun 2016.....	67
4.3	Peramalan Curah Hujan Pos Kedungadem.....	68
4.3.1	Peramalan Curah Hujan Menggunakan ARIMA.	68
4.3.2	Peramalan Curah Hujan Menggunakan SVR	84
4.3.3	Peramalan Curah Hujan Menggunakan GA-SVR	89
4.3.4	Perbandingan Model Terbaik Pos Kedungadem .	95
4.3.5	Kalender Tanam Padi Pos Kedungadem Periode Juli-Desember Tahun 2016.....	97
BAB V KESIMPULAN DAN SARAN		99
5.1	Kesimpulan	99
5.2	Saran	100
DAFTAR PUSTAKA		101
LAMPIRAN.....		107

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1 Tranformasi <i>Box-Cox</i>	8
Tabel 2.2 Karakteristik Teoritis Model ARIMA	10
Tabel 3.1 Struktur Data	25
Tabel 4.1 Statistika Deskriptif Data Curah Hujan Dasarian.....	33
Tabel 4.2 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak Tanpa <i>Differencing</i>	42
Tabel 4.3 <i>Diagnostic Checking</i> Model ARIMA Pos Cawak Tanpa <i>Differencing</i>	42
Tabel 4.4 Deteksi <i>Outlier</i> Model ARIMA Pos Cawak Tanpa <i>Differencing</i>	44
Tabel 4.5 Kriteria Keباikan Model ARIMA Pos Cawak Tanpa <i>Differencing</i>	45
Tabel 4.6 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak dengan <i>Differencing</i> Musiman 18 .	46
Tabel 4.7 <i>Diagnostic Checking</i> Model ARIMA Pos Cawak dengan <i>Differencing</i> Musiman 18 .	47
Tabel 4.8 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak dengan <i>Differencing</i> Musiman 36 .	49
Tabel 4.9 <i>Diagnostic Checking</i> Model ARIMA Pos Cawak dengan <i>Differencing</i> Musiman 36 .	50

Tabel 4.10	Deteksi <i>Outlier</i> Model ARIMA Pos Cawak dengan <i>Differencing</i> Musiman 36	52
Tabel 4.11	Hasil Uji Normalitas Residual Model ARIMA <i>Outlier</i> Pos Cawak dengan <i>Differencing</i> Musiman 36	53
Tabel 4.12	Kriteria Kebaikan Model ARIMA Pos Cawak dengan <i>Diferencing</i> Musiman 36..	53
Tabel 4.13	<i>Input Lag</i> SVR Pos Cawak	55
Tabel 4.14	Hasil Percobaan <i>Range</i> Parameter ϵ Pos Cawak	56
Tabel 4.15	Hasil Percobaan <i>Range</i> Parameter γ Pos Cawak	57
Tabel 4.16	Hasil Percobaan <i>Range</i> Parameter C Pos Cawak	57
Tabel 4.17	Kriteria Kebaikan Model SVR Pos Cawak	58
Tabel 4.18	Ilustrasi Nilai <i>Fitness</i> Setiap Kromosom Pos Cawak	60
Tabel 4.19	Ilustrasi Nilai <i>Fitness</i> Relatif dan <i>Fitness</i> Kumulatif Pos Cawak	61
Tabel 4.20	Ilustrasi Nilai Perbandingan <i>Fitness</i> Kumulatif dan Bilangan Random Pos Cawak	62
Tabel 4.21	Ilustrasi Proses <i>Elitism</i> Generasi Pertama Pos Cawak	64
Tabel 4.22	Ilustrasi Proses <i>Elitism</i> Generasi Kedua Pos Cawak	64

Tabel 4.23	Hasil Optimasi Parameter SVR Pos Cawak Menggunakan GA	65
Tabel 4.24	Hasil Perbandingan Metode Peramalan Pos Cawak	65
Tabel 4.25	Hasil <i>Forecast</i> Curah Hujan 6 Periode Selanjutnya Pos Cawak	66
Tabel 4.26	Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem Tanpa <i>Differencing</i>	72
Tabel 4.27	<i>Diagnostic Checking</i> Model ARIMA Pos Kedungadem Tanpa <i>Differencing</i>	74
Tabel 4.28	Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem dengan <i>Differencing</i> Musiman 18	76
Tabel 4.29	<i>Diagnostic Checking</i> Model ARIMA Pos Kedungadem dengan <i>Differencing</i> Musiman 18	77
Tabel 4.30	Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem dengan <i>Differencing</i> Musiman 36	79
Tabel 4.31	<i>Diagnostic Checking</i> Model ARIMA Pos Kedungadem dengan <i>Differencing</i> Musiman 36	80
Tabel 4.32	Deteksi <i>Outlier</i> Model ARIMA Pos Kedungadem dengan <i>Differencing</i> Musiman 36	81

Tabel 4.33	Hasil Uji Normalitas Residual Model ARIMA <i>Outlier</i> Pos Kedungadem dengan <i>Differencing</i> Musiman 36.....	82
Tabel 4.34	Kriteria Kebaikan Model ARIMA Pos Kedungadem dengan <i>Diferencing</i> Musiman 36	83
Tabel 4.35	<i>Input Lag</i> SVR Pos Kedungadem	84
Tabel 4.36	Hasil Percobaan <i>Range</i> Parameter ε Pos Kedungadem	86
Tabel 4.37	Hasil Percobaan <i>Range</i> Parameter γ Pos Kedungadem	86
Tabel 4.38	Hasil Percobaan <i>Range</i> Parameter C Pos Kedungadem	87
Tabel 4.39	Kriteria Kebaikan Model SVR Pos Kedungadem	87
Tabel 4.40	Ilustrasi Nilai <i>Fitness</i> Setiap Kromosom Pos Kedungadem	90
Tabel 4.41	Ilustrasi Nilai <i>Fitness</i> Relatif dan <i>Fitness</i> Kumulatif Pos Kedungadem	91
Tabel 4.42	Ilustrasi Nilai Perbandingan <i>Fitness</i> Kumulatif dan Bilangan Random Pos Kedungadem	91
Tabel 4.43	Ilustrasi Proses <i>Elitism</i> Generasi Pertama Pos Kedungadem	93
Tabel 4.44	Ilustrasi Proses <i>Elitism</i> Generasi Kedua Pos Kedungadem	94

Tabel 4.45	Hasil Optimasi Parameter SVR Pos Kedungadem Menggunakan GA	94
Tabel 4.46	Hasil Perbandingan Metode Peramalan Pos Kedungadem	95
Tabel 4.47	Hasil <i>Forecast</i> Curah Hujan 6 Periode Selanjutnya Pos Kedungadem	97

(Halaman Ini Sengaja Dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Kalender Tanam Padi	23
Gambar 3.1 Diagram Alir Penelitian	29
Gambar 3.2 Diagram Alir Peramalan Menggunakan ARIMA	30
Gambar 3.3 Diagram Alir Peramalan Menggunakan SVR	31
Gambar 3.4 Diagram Alir Peramalan Menggunakan GA-SVR.....	32
Gambar 4.1 <i>Boxplot</i> Curah Hujan Per Bulan Pos Cawak	34
Gambar 4.2 <i>Boxplot</i> Curah Hujan Per Tahun Pos Cawak	35
Gambar 4.3 <i>Boxplot</i> Curah Hujan Per Bulan Pos Kedungadem	36
Gambar 4.4 <i>Boxplot</i> Curah Hujan Per Tahun Pos Kedungadem	36
Gambar 4.5 <i>Time Series Plot</i> Curah Hujan Dasarian Pos Cawak.....	38
Gambar 4.6 Plot ACF Curah Hujan Dasarian Pos Cawak	38
Gambar 4.7 <i>Box-Cox</i> Data Curah Hujan Dasarian Pos Cawak	39
Gambar 4.8 <i>Box-Cox</i> Curah Hujan Dasarian Pos Cawak Hasil Transformasi.....	39
Gambar 4.9 <i>Time Series Plot</i> (a) dan Plot ACF Data Hasil Transformasi Pos Cawak	40

Gambar 4.10	Plot PACF Data Transformasi Curah Hujan Pos Cawak.....	41
Gambar 4.11	Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah <i>Differencing</i> Musiman 18 Pos Cawak	45
Gambar 4.12	Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah <i>Differencing</i> Musiman 36 Pos Cawak	49
Gambar 4.13	Ilustrasi Kromosom Parameter SVR Tiga Gen Untuk Pos Cawak.....	60
Gambar 4.14	Ilustrasi Kromosom yang Melakukan <i>Crossover</i> Pos Cawak	63
Gambar 4.15	Ilustrasi Proses Mutasi Pos Cawak	63
Gambar 4.16	Perbandingan Visual pada Data <i>Out Sample</i> Pos Cawak.....	66
Gambar 4.17	Kalender Tanam Padi Pos Cawak Periode Juli 2016 hingga Desember 2016.....	67
Gambar 4.18	<i>Time Series Plot</i> Curah Hujan Dasarian Pos Kedungadem	69
Gambar 4.19	Plot ACF Curah Hujan Dasarian Pos Kedungadem	69
Gambar 4.20	<i>Box-Cox</i> Data Curah Hujan Dasarian Pos Kedungadem	70
Gambar 4.21	<i>Box-Cox</i> Curah Hujan Dasarian Pos Kedungadem Hasil Transformasi.....	70
Gambar 4.22	<i>Time Series Plot</i> (a) dan Plot ACF Data Hasil Transformasi Pos Kedungadem.....	71
Gambar 4.23	Plot PACF Data Transformasi Curah Hujan Pos Kedungadem.....	72

Gambar 4.24	Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah <i>Differencing</i> Musiman 18 Pos Kedungadem	75
Gambar 4.25	Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah <i>Differencing</i> Musiman 36 Pos Kedungadem	78
Gambar 4.26	Ilustrasi Kromosom Parameter SVR Tiga Gen Untuk Pos Kedungadem.....	89
Gambar 4.27	Ilustrasi Kromosom yang Melakukan <i>Crossover</i> Pos Kedungadem	92
Gambar 4.28	Ilustrasi Proses Mutasi Pos Kedungadem ..	93
Gambar 4.29	Perbandingan Visual pada Data <i>Out Sample</i> Pos Kedungadem.....	96
Gambar 4.30	Kalender Tanam Padi Pos Kedungadem Periode Juli 2016 hingga Desember 2016..	97

(Halaman Ini Sengaja Dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Data Curah Hujan Dasarian di Pos Cawak Bulan Januari hingga Bulan Juni 2016	107
Lampiran 2 Data Curah Huja Dasarian di Pos Kedungadem Bulan Januari hingga Bulan Juni 2016.....	108
Lampiran 3 <i>Syntax</i> SAS ARIMA ([1,2,3,35],0,0) Pos Cawak	109
Lampiran 4 <i>Syntax</i> SAS ARIMA <i>Outlier</i> ([1,2,3,35],0,0) Pos Cawak	110
Lampiran 5 <i>Output</i> ARIMA <i>Outlier</i> ([1,2,3,35],0,0) Pos Cawak.....	111
Lampiran 6 <i>Syntax</i> SAS ARIMA (1,0,1)(0,1,1) ³⁶ Pos Cawak	112
Lampiran 7 <i>Syntax</i> SAS ARIMA <i>Outlier</i> (1,0,1)(0,1,1) ³⁶ Pos Cawak.....	113
Lampiran 8 <i>Output</i> ARIMA <i>Outlier</i> (1,0,1)(0,1,1) ³⁶ Pos Cawak.....	114
Lampiran 9 <i>Syntax</i> SAS ARIMA (0,0,1)(0,1,1) ³⁶ Pos Cawak	115
Lampiran 10 <i>Syntax</i> SAS ARIMA <i>Outlier</i> (0,0,1)(0,1,1) ³⁶ Pos Cawak.....	116
Lampiran 11 <i>Output</i> ARIMA <i>Outlier</i> (0,0,1)(0,1,1) ³⁶ Pos Cawak.....	117
Lampiran 12 <i>Syntax</i> SAS ARIMA (0,0,2)(0,1,1) ³⁶ Pos Cawak	118

Lampiran 13	<i>Syntax SAS ARIMA Outlier</i> (0,0,2)(0,1,1) ³⁶ Pos Cawak.....	119
Lampiran 14	<i>Output ARIMA Outlier</i> (0,0,2)(0,1,1) ³⁶ Pos Cawak.....	120
Lampiran 15	<i>Syntax SAS ARIMA</i> ([1,36,37],0,0)(0,1,0) ³⁶ Pos Cawak	121
Lampiran 16	<i>Syntax SAS ARIMA Outlier</i> ([1,36,37],0,0) (0,1,0) ³⁶ Pos Cawak	122
Lampiran 17	<i>Output ARIMA Outlier</i> ([1,36,37],0,0) (0,1,0) ³⁶ Pos Cawak	123
Lampiran 18	<i>Syntax SVR Model ARIMA</i> (3,0,0) (1,1,0) ¹⁸ Pos Cawak	124
Lampiran 19	Hasil Penguraian Model ARIMA Pos Cawak	125
Lampiran 20	<i>Syntax GA-SVR</i> Pos Cawak	128
Lampiran 21	<i>Syntax SAS ARIMA</i> (0,0,2) (0,1,1) ³⁶ Pos Kedungadem	129
Lampiran 22	<i>Syntax SAS ARIMA Outlier</i> (0,0,2) (0,1,1) ³⁶ Pos Kedungadem	130
Lampiran 23	<i>Output ARIMA Outlier</i> (0,0,2) (0,1,1) ³⁶ Pos Kedungadem	131
Lampiran 24	<i>Syntax SAS ARIMA</i> (1,0,1) (0,1,1) ³⁶ Pos Kedungadem	132
Lampiran 25	<i>Syntax SAS ARIMA Outlier</i> (1,0,1) (0,1,1) ³⁶ Pos Kedungadem	133
Lampiran 26	<i>Output ARIMA Outlier</i> (1,0,1) (0,1,1) ³⁶ Pos Kedungadem	134

Lampiran 27	<i>Syntax</i> SAS ARIMA (1,0,[1,52]) (0,1,1) ³⁶ Pos Kedungadem	135
Lampiran 28	<i>Syntax</i> SAS ARIMA <i>Outlier</i> (1,0,[1,52]) (0,1,1) ³⁶ Pos Kedungadem	136
Lampiran 29	<i>Output</i> ARIMA <i>Outlier</i> (1,0,[1,52]) (0,1,1) ³⁶ Pos Kedungadem	137
Lampiran 30	<i>Syntax</i> SVR Model ARIMA ([1,16,37],0,0) (3,1,0) ¹⁸ Pos Kedungadem ..	138
Lampiran 31	Hasil Penguraian Model ARIMA Pos Kedungadem	139
Lampiran 32	<i>Syntax</i> GA-SVR Pos Kedungadem..... ..	142
Lampiran 33	Data Curah Hujan Pos Cawak dan Pos Kedungadem dari BMKG Karangploso Malang	143

(Halaman Ini Sengaja Dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Sektor Pertanian mempunyai peranan penting dalam struktur perekonomian Jawa Timur. Sektor pertanian menduduki urutan ketiga dalam penyumbang PDRB Jawa Timur setelah industri pengolahan dan perdagangan. Salah satu bagian dari sektor pertanian adalah produksi tanaman padi. Provinsi Jawa Timur dikenal sebagai lumbung padi nasional karena Jawa Timur memberikan kontribusi produksi padi sebesar 17,50% dari total produksi nasional. Daerah penghasil padi terbanyak adalah Kabupaten Jember sebesar 978.373 ton, Kabupaten Lamongan sebesar 959.135 ton dan Kabupaten Bojonegoro sebesar 847.857 ton (BPS, 2015).

Kabupaten Bojonegoro sebagai salah satu daerah yang dilewati aliran sungai Bengawan Solo di bagian utara merupakan salah satu daerah dataran rendah subur di Jawa Timur dengan pertanian yang ekstensif (BPS, 2016). Sebanyak 33,31% lahan di Kabupaten Bojonegoro digunakan sebagai lahan sawah dan sebagian besar berada di sepanjang aliran sungai Bengawan Solo serta sekitar 51,80% dari lahan sawah tersebut adalah tipe sawah tadah hujan yang mengandalkan ketersediaan air hujan dalam produksi padi. Sekitar 48% sawah lainnya merupakan sawah teknis yang mengandalkan irigasi dari sungai Bengawan Solo dan waduk (BPS, 2014). Sektor pertanian di Bojonegoro menyumbang PDRB terbesar kedua setelah sektor pertambangan yaitu mencapai 18,13% dari total PDRB tahun 2015 (BPS, 2016).

Pada beberapa tahun terakhir produksi padi di Kabupaten Bojonegoro mengalami fluktuasi. Pada tahun 2010 hingga 2014, produksi padi berturut-turut sebesar 899.732 ton, 675.697 ton, 808.112 ton, 806.548 ton dan 847.857 ton (BPS, 2016). Salah satu penyebab fluktuasi produksi padi adalah iklim ekstrem. Iklim ekstrem dapat mengakibatkan terjadinya bencana kekeringan atau musim kemarau yang panjang serta adanya hujan ekstrem yang mengakibatkan bencana banjir. Dilansir dari viva.co.id tahun 2015

terjadi musim kemarau di Jawa Timur yang menyebabkan gagal panen, dimana Kabupaten Bojonegoro masuk dalam tiga daerah paling terdampak kekeringan, sebanyak 135 hektar lahan sawah mengalami gagal panen. Sementara, banjir mengakibatkan petani harus memanen padi lebih cepat yang berimbas pada harga padi yang lebih murah sehingga petani mengalami kerugian. Oleh karena itu, diperlukan antisipasi untuk mengurangi kerugian petani, salah satu antisipasinya berupa peramalan curah hujan. Hasil ramalan curah hujan yang akurat akan mengurangi kerugian akibat kegagalan panen

Kabupaten Bojonegoro memiliki 22 tempat yang digunakan sebagai pos pengamatan curah hujan. Pos pengamatan curah hujan tersebut tersebar di 15 kecamatan di Kabupaten Bojonegoro (BPS, 2016). Penelitian ini mengambil studi kasus di wilayah selatan dan timur Kabupaten Bojonegoro yaitu di pos pengamatan curah hujan Cawak dan Kedungadem. Pos pengamatan Cawak terletak di Kecamatan Kepohbaru, dimana luas lahan sawah terbesar di Kabupaten Bojonegoro terdapat di kecamatan Kepohbaru dengan tipe sawah irigasi sebesar 3219 hektar dan non irigasi sebesar 3257 hektar. Kedungadem adalah salah satu kecamatan di Bojonegoro Selatan dengan tipe sawah dominan adalah sawah tadah hujan yang mengandalkan ketersediaan air dalam produksi padi (BPS, 2016). Selama ini untuk mencukupi pengairan per-sawah di wilayah timur dan selatan Bojonegoro mengandalkan cadangan air Waduk Pacal yang kini mulai menipis. Sebaliknya pada musim hujan, aliran air dari Bengawan Solo terlalu melimpah yang mengacaukan pola tanam. Oleh karena itu, keberhasilan produksi padi sangat bergantung pada informasi ketersediaan air melalui data curah hujan yang tersusun dalam kalender tanam.

Peramalan curah hujan menggunakan metode ARIMA telah banyak dilakukan. Namun, penggunaan metode ARIMA pada peramalan curah hujan seringkali belum mampu memberikan residual yang memenuhi asumsi berdistribusi normal serta belum mampu mengakomodasi adanya data *outlier*. Munarsih (2011) dalam Wiyanti (2012) menyatakan bahwa ARIMA sangat baik kete-

patannya untuk peramalan jangka pendek dan untuk data *time series non stasioner* pada saat linier. Untuk peramalan jangka panjang ketepatannya kurang baik serta mengalami penurunan keakuratan apabila terdapat komponen nonlinier *time series* pada data pengamatan. Zhang (2003) pun menyatakan bahwa ARIMA tidak mampu memodelkan *time series* yang nonlinier. Padahal data curah hujan seringkali menghasilkan data yang non linier. Oleh karena itu dibutuhkan metode peramalan yang dapat mengatasi kasus nonlinieritas pada data.

Salah satu metode peramalan yang dapat mengatasi kasus nonlinieritas adalah metode *Support Vector Regression* (SVR) dengan menambahkan fungsi kernel (Dimyati, 2014). Pada penelitian ini akan digunakan fungsi kernel *Radial Basis Function* (RBF). RBF merupakan fungsi kernel yang sering digunakan dan menghasilkan hasil yang lebih bagus. Lippi, Bertini dan Frasconi (2013) dan Harafani (2015) sebelumnya telah membuktikan bahwa fungsi kernel RBF menghasilkan nilai akurasi peramalan yang lebih baik dibandingkan fungsi kernel lain. Keunggulan lain metode SVR menurut Gunn (1998) adalah SVR dapat mengatasi kasus *overfitting* berdasarkan pada konsep *risk minimization*. Berbeda dengan metode ARIMA yang sulit dalam memenuhi asumsi, metode SVR adalah salah satu metode yang tidak memerlukan asumsi. Metode SVR juga dapat memberikan hasil yang baik walaupun hanya dengan data yang sedikit (Zhao, Tao, dan Zio, 2015). Pada metode SVR terdapat beberapa parameter yaitu *Cost*, ϵ dan γ , dimana nilai parameter ini harus diatur agar mendapatkan nilai parameter yang optimal. Salah satu permasalahan yang sering dialami ketika menggunakan SVR adalah penentuan parameter model yang optimal. Padahal, pemilihan parameter yang optimal mempunyai peranan penting untuk akurasi peramalan, sehingga pada penelitian ini akan digunakan *Genetic Algorithm* (GA) untuk optimasi nilai parameter SVR dengan harapan dapat menghasilkan nilai akurasi yang lebih bagus. Makridakis dan Hibon (2000) menyatakan bahwa kelebihan dari menggabungkan model menjadi satu adalah menghasilkan ramalan dengan tingkat akurasi yang lebih baik secara rata-rata di-

bandingkan dengan model tunggal lainnya. *Genetic Algorithm* (GA) memiliki beberapa keunggulan dalam proses optimasi yaitu GA bekerja pada kumpulan solusi dan GA mencari berdasarkan populasi dari solusi, bukan hanya satu solusi (Sawaka, 2002 diacu dalam Putri, 2016).

Penelitian dengan menggunakan metode SVR pernah dilakukan oleh Msiza, Nelwamondo dan Marwala (2008) untuk meramalkan permintaan air dengan hasil penelitian yaitu hasil prediksi yang dihasilkan oleh ANN dan SVR memberikan hasil error yang lebih kecil dibandingkan metode statistika biasa. Penelitian metode GA-SVR pernah dilakukan oleh Gu, Zhu dan Jiang (2011) untuk meramalkan harga perabotan rumah dimana metode GA-SVM menghasilkan peramalan yang lebih baik dibandingkan menggunakan *grid algorithm*. Chen K. Y (2007) juga menggunakan GA-SVR untuk meramalkan *system realibility* dengan hasil penelitian yaitu metode GA-SVR menghasilkan nilai MAPE dan N-RMSE yang lebih kecil dibandingkan dengan metode NN dan ARIMA. Guo, Li dan Zhang tahun 2012 membandingkan GA-SVR dan SVR tanpa menggunakan GA untuk meramalkan harga minyak dimana menggunakan GA-SVR dihasilkan peramalan yang lebih baik dibandingkan SVR tanpa GA.

Penelitian kali ini akan menerapkan metode ARIMA, SVR dan GA-SVR untuk meramalkan curah hujan di Kabupaten Bojonegoro. Metode ARIMA tetap digunakan untuk perbandingan dengan alasan bahwa metode yang kompleks belum tentu memberikan hasil peramalan yang lebih akurat (Makridakis dan Hibon, 2000). Kriteria perbandingan akurasi ramalan yang digunakan adalah RMSE dan SMAPE. Berdasarkan pada hasil *forecast* bulan selanjutnya akan disusun kalender tanam padi dengan harapan dapat memberikan tambahan informasi bagi petani dalam memilih waktu yang tepat untuk melakukan penanaman padi pada tipe sawah tadah hujan.

1.2 Rumusan Masalah

Kabupaten Bojonegoro merupakan tiga kabupaten penghasil padi terbesar di Provinsi Jawa Timur. Pada tahun 2010-2014 pro-

duksi padi Bojonegoro mengalami fluktuasi. Salah satu penyebab fluktuasi produksi padi adalah iklim ekstrem. Iklim ekstrem mengakibatkan musim kemarau yang panjang dan hujan ekstrem. Pada tahun 2015 Kabupaten Bojonegoro termasuk 3 daerah yang paling terdampak kekeringan di Jawa Timur. Sebaliknya pada musim hujan Kabupaten Bojonegoro sering mengalami bencana banjir yang mengakibatkan petani harus memanen padi lebih awal. Oleh karena itu, dibutuhkan antisipasi untuk mengurangi kerugian petani salah satu antisipasinya berupa peramalan curah hujan yang akurat. Beberapa metode peramalan yang dapat digunakan pada curah hujan adalah ARIMA dan *Support Vector Regression* (SVR). Metode SVR digunakan dikarenakan metode ARIMA biasanya baik pada peramalan data yang linier, sedangkan metode SVR bisa mengatasi kasus non linieritas yang merupakan karakteristik data curah hujan dengan menambahkan fungsi kernel. Untuk mengoptimalkan parameter model SVR maka digunakan *Genetic Algorithm* (GA) dengan harapan dapat meningkatkan akurasi peramalan. Berdasarkan permasalahan yang telah diuraikan, maka dapat dirumuskan suatu permasalahan, yaitu meramalkan curah hujan di Kabupaten Bojonegoro pada Pos Cawak dan Pos Kedungadem menggunakan metode ARIMA, SVR dan GA-SVR. Dari ketiga metode tersebut akan dibandingkan berdasarkan kriteria akurasi RMSE dan SMAPE. Metode terbaik akan digunakan untuk meramalkan curah hujan di Kabupaten Bojonegoro 6 bulan selanjutnya. Dari hasil peramalan 6 bulan selanjutnya, akan digunakan untuk menyusun kalender tanam padi untuk tipe sawah tadah hujan pada Pos Cawak dan Pos Kedungadem.

1.3 Tujuan

Berdasarkan uraian rumusan masalah dapat diambil tujuan dari penelitian ini sebagai berikut,

1. Mendeskripsikan karakteristik curah hujan di Kabupaten Bojonegoro.
2. Meramalkan curah hujan di Kabupaten Bojonegoro menggunakan metode ARIMA, SVR dan GA-SVR.

3. Memperoleh metode terbaik dari tiga metode yaitu ARIMA, SVR dan GA-SVR untuk melakukan peramalan curah hujan di Kabupaten Bojonegoro.
4. Memperoleh hasil nilai ramalan curah hujan di Kecamatan Kabupaten Bojonegoro berdasarkan metode terbaik.
5. Membuat kalender tanam padi berdasarkan hasil peramalan untuk tipe sawah tadah hujan.

1.4 Manfaat

Dari penelitian ini, manfaat yang dapat diambil adalah sebagai berikut,

1. Memberikan tambahan informasi bagi :
 - a. BMKG dalam memprediksikan curah hujan di Kabupaten Bojonegoro di masa mendatang.
 - b. Dinas Pertanian Kabupaten Bojonegoro dalam menyusun kalender tanam padi untuk sawah tadah hujan pada Pos Cawak dan Kedungadem yang dapat digunakan sebagai penentuan waktu tanam padi yang tepat.
2. Memberikan Informasi kepada masyarakat tentang keadaan curah hujan di Kabupaten Bojonegoro.

1.5 Batasan Masalah

Batasan penelitian masalah yang digunakan pada penelitian ini adalah data yang digunakan merupakan data curah hujan harian di Kabupaten Bojonegoro tahun 2000-2016 pada pos pengamatan yaitu Pos Cawak dan Pos Kedungadem. Metode yang digunakan adalah ARIMA, SVR dan GA-SVR. Metode SVR dipilih karena dapat mengatasi kasus non linieritas pada data curah hujan serta metode yang tidak memerlukan asumsi. Untuk optimasi parameter SVR digunakan *Genetic Algorithm* (GA) dengan harapan untuk meningkatkan akurasi peramalan. Metode linier ARIMA tetap digunakan untuk perbandingan dikarenakan metode yang kompleks belum tentu menghasilkan peramalan yang lebih akurat.

BAB II

TINJAUAN PUSTAKA

2.1 *Autoregressive Integrated Moving Avarage (ARIMA)*

ARIMA merupakan suatu metode peramalan yang biasanya sangat baik digunakan untuk melakukan peramalan jangka pendek. Selain itu model ARIMA juga merupakan salah satu model yang digunakan dalam peramalan data *time series* yang bersifat non stasioner. Secara umum model ARIMA (p, d, q) atau ARIMA *non seasonal* ditulis sebagai berikut (Wei, 2006).

$$\phi_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B)a_t \quad (2.1)$$

dimana,

p : orde dari AR

q : orde dari MA

$\phi_p(B)$: koefisien komponen AR orde p

$(1-B)^d$: operator untuk differencing orde d

a_t : residual yang *white noise* dan berdistribusi normal

2.1.1 Stasioneritas Model ARIMA

Suatu data *time series* Y_t bersifat stasioner dalam *mean* dan varians, maka *mean* dan varians tidak dipengaruhi oleh waktu pengamatan, dengan demikian,

Mean dari Y_t :

$$E(Y_t) = E(Y_{t+k}) = \mu \quad (2.2)$$

Varians dari Y_t :

$$E(Y_t - \mu)^2 = E(Y_{t+k} - \mu)^2 = \sigma^2 \quad (2.3)$$

Auto kovarians merupakan kovarians antara Y_t dan Y_{t+k} .

$$Cov(Y_t, Y_{t+k}) = E[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k \quad (2.4)$$

Pada sembarang nilai t dan k , dimana k adalah *time lag*.

Stasioner dibagi menjadi dua, yaitu *strictly stationary* dan *weakly stationary*. Proses stokastik Y_t dikatakan *strictly stationary* apabila distribusi bersama dari $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ sama dengan distribusi bersama dari $Y_{t_1-k}, Y_{t_2-k}, \dots, Y_{t_n-k}$ untuk semua titik waktu t_1, t_2, \dots, t_n dan semua pilihan *time lag* k . Sementara itu, proses stokastik Y_t dikatakan *weakly stasionary* jika fungsi

rata-rata konstan pada setiap waktu dan $\gamma_{t,t-k} = \gamma_{0,k}$ untuk semua waktu t dan *time lag* k , di mana $\gamma_{t,t-k}$ adalah autokovarians antara Y_t dengan Y_{t-k} (Cryer dan Chan, 2008).

Suatu deret waktu Y_t dikatakan tidak stasioner terhadap varians, jika Y_t berubah sejalan dengan perubahan level varians dimana c merupakan konstanta. *Box* dan *Cox* memberikan suatu ide transformasi terhadap varians yang tidak konstan dengan menggunakan *power transformation* sebagai berikut (Wei, 2006).

$$T(Y_t) = \frac{Y_t^{\lambda-1}}{\lambda}, \text{ dimana } \lambda \neq 0 \quad (2.5)$$

λ adalah parameter transformasi, untuk $\lambda = 0$ dilakukan pendekatan sebagai berikut:

$$\lim_{\lambda \rightarrow 0} T(Y_t) = \lim_{\lambda \rightarrow 0} (Y_t)^\lambda = \lim_{\lambda \rightarrow 0} \frac{Y_t^{\lambda-1}}{\lambda} = \ln(Y_t) \quad (2.6)$$

Secara umum nilai λ dan pilihan transformasi yang digunakan sebagai berikut (Wei, 2006).

Tabel 2. 1 Transformasi *Box-cox*

Estimasi λ	Transformasi
-1,0	$\frac{1}{Y_t}$
-0,5	$\frac{1}{\sqrt{Y_t}}$
0	$\ln(Y_t)$
0,5	$\sqrt{Y_t}$
1	Y_t (tidak ada transformasi)

Deret waktu dikatakan tidak stasioner dalam *mean*, jika plot data terhadap sumbu waktu (t) membentuk *trend* tertentu terhadap garis sumbu waktu (t). Untuk memastikan adanya suatu *trend* stokastik pada data *series* Y_t , digunakan indikator sebagai berikut (Makridakis, Wheelwright, dan McGee, 1999).

1. Pola *Autocorrelation function* (ACF) suatu *time series* (Y_t) menurun secara lambat, sedangkan pola *Partial Autocorrelation function* (PACF) *cut off* pada lag 1.
2. *Mean* level yang berubah dalam arah tertentu.

3. Parameter model dari *series* Y_t , tidak memenuhi syarat kestasioneran.

Selain secara visual, kestasioneran data dalam *mean* dapat dideteksi dengan melakukan pengujian *Augmented Dickey Fuller* (ADF). Uji ADF merupakan pengembangan uji *Dickey Fuller* (DF) yang digunakan untuk menguji kestasioneran data dalam *mean*. Uji ADF mengakomodasi terjadinya korelasi pada residual dengan menambahkan *lag-lag* dari variabel depende Y_t (Gujarati, 2004). Uji ADF secara spesifik mengikuti estimasi regresi sebagai berikut.

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \varepsilon_t \quad (2.7)$$

Berikut adalah hipotesis dari uji ADF.

$H_0 : \delta = 0$ (data tidak stasioner terhadap *mean*)

$H_1 : \delta < 0$ (data stasioner terhadap *mean*)

Statistik uji :

$$T = \frac{\hat{\delta}}{SE(\hat{\delta})} \quad (2.8)$$

Daerah kritis : tolak H_0 apabila nilai T lebih besar dari titik kritis T^* pada tabel *Dickey Fuller*.

Cara yang digunakan untuk mengatasi kondisi non-stasioner dalam *mean* adalah dengan melakukan pembedaan (*differencing*) terhadap data dengan persamaan sebagai berikut (Cryer dan Chan, 2008).

$$W_t = Y_t - Y_{t-1} \quad (2.9)$$

dimana W_t merupakan nilai *series* Y_t setelah dilakukan *differencing*. *Differencing* dapat dinyatakan sebagai berikut:

$$W_t = (1 - B)^d Y_t \quad (2.10)$$

keterangan:

d : orde *differencing*.

Y_t : nilai observasi pada waktu ke- t .

2.1.2 Identifikasi Model ARIMA

Dalam identifikasi model ARIMA dilakukan dengan melihat plot ACF dan PACF. Fungsi *autocorrelation function* (ACF) merupakan suatu hubungan linier antara pengamatan Y_t dengan pengamatan Y_{t-k} (Cryer dan Chan, 2008).

$$\hat{\rho}_k = \hat{r}_k = \frac{\text{Cov}(Y_t, Y_{t-k})}{\sqrt{\text{var } Y_t} \sqrt{\text{var } Y_{t-k}}}, k = 1, 2, \dots \quad (2.11)$$

dimana jika data diasumsikan stasioner terhadap *mean* dan varians maka diperoleh rumus ACF sebagai berikut.

$$\hat{\rho}_k = \hat{r}_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}, k = 1, 2, \dots \quad (2.12)$$

Fungsi *partial autocorrelation function* (PACF) digunakan untuk menunjukkan besarnya hubungan antar nilai variabel yang sama, dengan menganggap pengaruh dari semua kelambatan waktu yang lain adalah konstan (Cryer dan Chan, 2008).

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j} \quad (2.13)$$

dimana,

$$\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j} \text{ untuk } j=1, 2, \dots, k-1$$

Pada tahap identifikasi model ARIMA dapat dilakukan dengan melihat plot dari ACF dan PACF. Beberapa model ARIMA dari plot ACF dan PACF adalah sebagai berikut (Wei, 2006).

Tabel 2. 2 Karakteristik Teoritis Model ARIMA

Model	ACF	PACF
AR (p)	<i>Dies Down</i>	<i>Cut off</i> setelah lag p
MA (q)	<i>Cut off</i> setelah lag q	<i>Dies Down</i>
ARMA (p, q)	<i>Dies Down</i>	<i>Dies Down</i>
AR (p) atau MA (q)	<i>Cut off</i> setelah lag q	<i>Cut off</i> setelah lag p

2.1.3 Estimasi Parameter Model ARIMA

Metode estimasi yang digunakan dalam penelitian ini adalah metode *least square* (*Conditional Least Square*). Metode ini bekerja dengan meminimumkan jumlah kuadrat error (SSE) dalam model. (Cryer dan Chan, 2008). Misalkan model AR(p):

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \dots + \phi_p(Y_{t-p} - \mu) + a_t \quad (2.14)$$

Model AR(p) tersebut dapat dilihat sebagai model regresi dengan $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ sebagai variabel prediktor dan Y_t sebagai variabel respon. Metode *least square* meminimumkan jumlah kuadrat error dari differensiasi :

$$Y_t - \mu - \phi_1(Y_{t-1} - \mu) - \phi_2(Y_{t-2} - \mu) - \dots - \phi_p(Y_{t-p} - \mu) \quad (2.15)$$

Berdasarkan persamaan tersebut, diketahui bahwa Y_1, Y_2, \dots, Y_n diamati, maka proses penjumlahan dapat dilakukan menggunakan $t = p + 1, p + 2, \dots, n$, sehingga metode estimasi *least square* meminimumkan persamaan berikut ini.

$$S_c(\phi_1, \dots, \phi_p, \mu) = \sum_{t=p+1}^n \left[Y_t - \mu - \phi_1(Y_{t-1} - \mu) - \dots - \phi_p(Y_{t-p} - \mu) \right]^2 \quad (2.16)$$

Persamaan (2.16) disebut juga fungsi *conditional sum of squares*. Berdasarkan prinsip dari *least squares*, parameter ϕ dan μ di estimasi menggunakan masing-masing nilai yang meminimumkan $S(\phi, \mu)$. Adapun persamaan (2.17) merupakan proses penurunan $\frac{\partial S_c}{\partial \mu} = 0$.

$$\sum_{t=p+1}^n 2 \left[Y_t - \mu - \phi_1(Y_{t-1} - \mu) - \dots - \phi_p(Y_{t-p} - \mu) \right] (-1 + \phi_1 + \dots + \phi_p) = 0 \quad (2.17)$$

Hasil penurunan yang ditunjukkan oleh persamaan (2.17) memberikan nilai estimasi parameter μ dari model AR (p) hingga nilai estimasi seperti yang ditunjukkan oleh persamaan (2.18) berikut.

$$\mu = \frac{1}{(n-p)(-1 + \phi_1 + \dots + \phi_p)} \left[\sum_{t=p+1}^n Y_t - \phi_1 \sum_{t=p+1}^n Y_{t-1} - \dots - \phi_p \sum_{t=p+1}^n Y_{t-p} \right] \quad (2.18)$$

Untuk n berukuran besar diketahui bahwa:

$$\frac{1}{n-p} \sum_{t=p+1}^n Y_t \approx \frac{1}{n-p} Y_{t-1} \approx \frac{1}{n-p} Y_{t-2} \approx \dots \approx \frac{1}{n-p} Y_{t-p} \approx \bar{Y} \quad (2.19)$$

Berdasarkan persamaan (2.19) diketahui bahwa nilai mendekati \bar{Y} sehingga persamaan (2.18) dapat disederhanakan menjadi sebagai berikut.

$$\mu = \frac{1}{(n-p)(-1 + \phi_1 + \dots + \phi_p)} \left[\bar{Y} - \phi_1 \bar{Y} - \phi_2 \bar{Y} - \dots - \phi_p \bar{Y} \right] = \bar{Y} \quad (2.20)$$

Selanjutnya, *least square* meminimumkan $S_c(\phi_1, \phi_2, \dots, \phi_p)$ untuk mendapatkan estimasi parameter dari ϕ_i . Adapun estimasi

parameter ϕ_1 didapatkan dari $\frac{\partial S_c}{\partial \phi_1} = 0$ yang ditunjukkan oleh persamaan berikut ini.

$$\frac{\partial S_c}{\partial \phi_1} = -2 \sum_{t=p+1}^n \left[(Y_t - \bar{Y}) - \phi_1 (Y_{t-1} - \bar{Y}) - \dots - \phi_p (Y_t - \bar{Y}) \right] (Y_{t-1} - \bar{Y}) = 0 \quad (2.21)$$

Persamaan (2.21) dapat dituliskan menjadi persamaan berikut.

$$\begin{aligned} \sum_{t=p+1}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y}) &= \left[\sum_{t=p+1}^n (Y_{t-1} - \bar{Y})^2 \right] \phi_1 + \left[\sum_{t=p+1}^n (Y_{t-1} - \bar{Y})(Y_{t-2} - \bar{Y}) \right] \phi_2 \\ &+ \dots + \left[\sum_{t=p+1}^n (Y_{t-1} - \bar{Y})(Y_{t-p} - \bar{Y}) \right] \phi_p \end{aligned} \quad (2.22)$$

Penjumlahan dari hasil $\sum_{t=p+1}^n (Y_t - \bar{Y})(Y_{t-1} - \bar{Y})$ mendekati pembilang r_1 . Namun, tidak terdapat hasil $(Y_p - \bar{Y})(Y_{p-1} - \bar{Y})$. Kondisi yang sama juga terdapat pada persamaan $\sum_{t=p+1}^n (Y_{t-1} - \bar{Y})(Y_{t-p} - \bar{Y})$ yang tidak menyertakan suatu hasil $(Y_n - \bar{Y})(Y_{n-1} - \bar{Y})$. Apabila kedua ruas pada persamaan (2.22) d-ibagi dengan $\sum_{t=p+1}^n (Y_t - \bar{Y})^2$, maka hasil yang diperoleh sebagai berikut.

$$r_1 = \phi_1 + r_1 \phi_2 + r_2 \phi_3 + \dots + r_{p-1} \phi_p \quad (2.23)$$

Estimasi untuk parameter $\phi_2, \phi_3, \dots, \phi_p$ dapat dilakukan dengan cara yang sama sedemikian hingga hasil estimasi adalah dengan menyelesaikan persamaan berikut ini.

$$\left. \begin{aligned} \phi_1 + r_1 \phi_2 + r_2 \phi_3 + \dots + r_{p-1} \phi_p &= r_1 \\ r_1 \phi_1 + \phi_2 + r_1 \phi_3 + \dots + r_{p-2} \phi_p &= r_2 \\ &\vdots \\ r_{p-1} \phi_1 + r_{p-2} \phi_2 + r_{p-3} \phi_3 + \dots + \phi_p &= r_p \end{aligned} \right\} \quad (2.24)$$

Persamaan (2.24) disebut juga persamaan Yule-Walker untuk model AR(p). Berdasarkan persamaan Yule-Walker r_k merupakan pengganti dari ρ_k , $k = 1, 2, \dots, p$ dengan ρ_k merupakan autokorelasi lag ke- k .

Misalkan diketahui model ARMA (1,1)

$$Y_t - \phi_1 Y_{t-1} = a_t - \theta_1 a_{t-1} \quad (2.25)$$

Untuk menghitung a_t maka

$$\begin{aligned} a_t &= Y_t - \phi_1 Y_{t-1} + \theta_1 a_{t-1} \\ a_t &= Y_t - \phi_1 Y_{t-1} + \theta_1 (Y_{t-1} - \phi_1 Y_{t-2} + \theta_1 a_{t-2}) \quad (2.26) \\ a_t &= Y_t - (\phi_1 - \theta_1) Y_{t-1} - \phi_1 \theta_1 Y_{t-2} + \theta_1^2 a_{t-2} \end{aligned}$$

:

Persamaan (2.26) adalah persamaan yang tidak linier di parameter, sehingga untuk model ARMA maka prosedur estimasi non linier *least square* akan digunakan untuk mengestimasi parameter.

Prosedur non linier *least square* akan dilakukan teknik iterasi, karena model linier adalah kasus khusus dari model non linier. Berikut adalah persamaan utama dari *non linier least square* dengan menggunakan model regresi linier.

$$\begin{aligned} Y_t &= E(Y_t | X_{t_i}'s) + a_t \\ &= \alpha_1 X_{t_1} + \alpha_2 X_{t_2} + \dots + \alpha_p X_{t_p} + a_t \end{aligned} \quad (2.27)$$

dimana $t = 1, 2, \dots, n$ dan $a_t's$ adalah i.i.d $N(0, \sigma^2_a)$ independen untuk setiap X_{t_i} . Dari persamaan sebelumnya, estimasi *least square* (linier atau non linier) biasa menggunakan perhitungan dengan teknik iterasi sebagai berikut.

- a. Menetapkan vektor inisialisasi yaitu $\tilde{\alpha}$, kemudian menghitung residual $\tilde{e} = (Y - \tilde{Y})$ dan residual dari *sum square*.

$$S(\tilde{\alpha}) = \tilde{e}'\tilde{e} = (Y - \tilde{Y})'(Y - \tilde{Y}) \quad (2.28)$$

dimana $\tilde{Y} = f(\tilde{\alpha})$ adalah vektor prediksi dengan mengganti parameter yang tidak diketahui dengan suatu nilai hasil prediksi. Model $f(X_t, \alpha)$ didekati dengan persamaan deret Taylor orde pertama dengan mengekspansi nilai inisial dari $\tilde{\alpha}$ sebagai berikut.

$$f(\alpha) = f(\tilde{\alpha}) + \bar{X}_{\tilde{\alpha}}\delta \quad (2.29)$$

dimana $\delta = (\alpha - \tilde{\alpha})$ dan $\bar{X}_{\tilde{\alpha}} = [X_{ij}]$ adalah matriks $n \times p$ yang merupakan turunan partial dari $\tilde{\alpha}$. Kemudian dihitung

$$\delta = (\bar{X}_{\tilde{\alpha}}' \bar{X}_{\tilde{\alpha}})^{-1} \bar{X}_{\tilde{\alpha}}' \tilde{e} = (\delta_1, \delta_1, \dots, \delta_p)' \quad (2.30)$$

Untuk non linier model maka $\bar{X}_{\tilde{\alpha}}$ akan berubah pada setiap iterasi.

- b. Menetapkan estimasi *least square* terbaru

$$\hat{\alpha} = \tilde{\alpha} + \delta \quad (2.31)$$

dan residual dari *sum square* $S(\hat{\alpha})$. Untuk model non linier langkah b ini hanya akan menghasilkan nilai insial baru untuk digunakan pada iterasi selanjutnya.

Jika diberikan model ARMA (p,q) , maka untuk menyelesaikannya bisa digunakan prosedur non linier *least square* untuk menghitung estimasi *least square* dimana meminimalkan eror dari *sum square* $S_*(\phi, \mu, \theta)$ atau $S(\phi, \mu, \theta)$. Iterasi akan terus dilakukan sampai kriteria konvergen tercapai. Kriteria konvergen adalah reduksi relatif dari *sum square*, dimana maksimum selisih nilai parameter dari setiap iterasi kurang dari suatu nilai yang ditetapkan atau jumlah iterasi lebih besar dari nilai yang ditetapkan (Wei, 2006).

2.1.4 Pengujian Signifikansi Parameter

Model ARIMA yang baik dan yang dapat menggambarkan suatu kejadian adalah model yang salah satunya menunjukkan bahwa estimasi parameternya berbeda signifikan dengan nol. Setelah didapatkan nilai estimasi dari parameter, maka selanjutnya menguji signifikansi parameter tersebut. Jika $i=1,2,\dots, p$, maka hipotesis yang digunakan untuk melakukan pengujian signifikansi parameter model AR adalah sebagai berikut. (Bowerman, O'Connell & Koehler, 2004).

$H_0 : \phi_i = 0$ (Parameter AR bernilai sama dengan nol atau tidak signifikan)

$H_1 : \phi_i \neq 0$ (Parameter AR bernilai tidak sama dengan nol atau signifikan)

Statistik uji :

$$t_{hitung,i} = \frac{\hat{\phi}_i}{SE(\hat{\phi}_i)} \quad (2.32)$$

H_0 akan ditolak apabila nilai statistik uji $|t_{hitung,i}| > t_{\alpha/2, (n-n_p)}$ atau nilai $p\text{-value} < \alpha$ dimana n_p adalah banyaknya parameter AR pada model, yaitu $n_p=p+1$ apabila ada *intercept* dan $n_p=p$ apabila tidak ada *intercept*. Sedangkan jika $j=1,2,\dots,q$, hipotesis yang digunakan untuk melakukan pengujian signifikansi model MA adalah :

$H_0 : \theta_j = 0$ (Parameter MA bernilai sama dengan nol atau tidak signifikan)

$H_1 : \theta_j \neq 0$ (Parameter MA bernilai tidak sama dengan nol atau signifikan)

Statistik uji :

$$t_{hitung,j} = \frac{\hat{\theta}_j}{SE(\hat{\theta}_j)} \quad (2.33)$$

H_0 akan ditolak apabila nilai statistik uji $|t_{hitung,j}| > t_{\alpha/2, (n-n_q)}$ atau nilai $p\text{-value} < \alpha$ dimana n_q adalah banyaknya parameter MA pada model, yaitu $n_q=q+1$ apabila ada *intercept* dan $n_q=q$ apabila tidak ada *intercept*.

2.1.5 Diagnostic Checking

1. Uji Asumsi Residual *White Noise*

Pada model ARIMA, residual yang diperoleh harus memenuhi asumsi yaitu *white noise*. Pengujian *white noise* menggunakan hipotesis sebagai berikut (Wei, 2006).

$H_0 : \rho_1 = \rho_2 = \dots = \rho_K = 0$ (residual memenuhi syarat *white noise*).

H_1 : Minimal ada satu $\rho_k \neq 0$ dengan $k=1,2,\dots, K$ (residual tidak memenuhi syarat *white noise*).

Statistik Uji :

$$Q = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{(n-k)} \quad (2.34)$$

Daerah penolakan : tolak H_0 jika $Q > \chi^2_{\alpha, k-p-q}$ atau $P\text{-value} < \alpha$

2. Uji Asumsi Normalitas Residual

Uji asumsi normalitas residual, dapat dilakukan dengan uji Kolmogorov-smirnov sebagai berikut (Daniel, 1989).

$H_0 : F(x) = F_0(x)$ (residual berdistribusi normal)

$H_1 : F(x) \neq F_0(x)$ (residual tidak berdistribusi normal)

Statistik Uji

$$D = \sup_x |S(x) - F_0(x)| \quad (2.35)$$

Daerah penolakan : tolak H_0 jika $D > D_{(1-\alpha, n)}$ atau $P\text{-value} < \alpha$.

2.2 Uji Nonlinieritas Terasvirta

Uji nonlinieritas dilakukan untuk mengetahui apakah data mengikuti pola linier atau nonlinier. Salah satu uji nonlinieritas yaitu uji Terasvirta. Terasvirta, Lin dan Granger (1993) memperkenalkan suatu uji nonlinieritas yang termasuk uji kelompok

Lagrange Multiplier (LM) yang dikembangkan dari model *neural network*. Berikut adalah hipotesis pada uji Terasvirta.

H_0 : $f(x)$ adalah fungsi linier dalam x atau model linier

H_1 : $f(x)$ adalah fungsi non linier dalam x atau model non linier

Prosedur untuk mendapatkan nilai statistik uji F adalah sebagai berikut.

1. Meregresikan Y_t dengan 1, Y_{t-1} , Y_{t-2} , ..., Y_{t-k} dan menghitung residual \hat{a}_t serta menghitung jumlah kuadrat residual.

$$SSR_0 = \sum_{t=1}^T \hat{a}_t^2 \quad (2.36)$$

2. Meregresikan \hat{a}_t dengan 1, Y_{t-1} , Y_{t-2} , ..., Y_{t-k} dan m prediktor tambahan yang merupakan suku kubik atau kuadrat yang merupakan hasil pendekatan ekspansi deret Taylor. Menghitung residual \hat{v}_t serta menghitung jumlah kuadrat residual.

$$SSR = \sum_{t=1}^T \hat{v}_t^2 \quad (2.37)$$

3. Menghitung nilai

$$F = \frac{(SSR_0 - SSR)/m}{SSR/(n-p-1-m)} \quad (2.38)$$

dengan n adalah jumlah pengamatan.

Dibawah hipotesis linieritas dalam mean, nilai F didekati dengan distribusi F dengan derajat bebas m dan $n - p - 1 - m$ (Terasvirta, 1993 diacu dalam Putri, 2016).

2.3 Support Vector Regression (SVR)

SVR merupakan bagian dari *Support Vector Machine* (SVM) yang diperkenalkan oleh Vapnik tahun 1995. SVM adalah sistem pembelajaran yang menggunakan ruang hipotesis berupa fungsi-fungsi linier dalam sebuah ruang fitur (*feature space*) berdimensi tinggi, dilatih dengan algoritma pembelajaran yang didasarkan pada teori optimasi dengan mengimplementasikan *learning* bias. Konsep SVM menggunakan konsep *ϵ -insensitive loss function*. SVM dapat digeneralisasi untuk melakukan pendekatan fungsi yang dikenal dengan SVR. SVR merupakan metode pengembangan dari *Support Vector Machine* untuk kasus prediksi dan regresi (Gunn, 1998). Seperti halnya SVM, SVR bertujuan untuk menciptakan *hyperplane* dan *support vector* untuk menentukan nilai prediksi. SVR memiliki performa yang handal dalam prediksi

data *time-series*. SVR merupakan metode yang dapat mengatasi *overfitting*, sehingga menghasilkan kinerja yang bagus. Jika nilai $\varepsilon = 0$, maka didapat suatu regresi yang sempurna (Santosa, 2007 diacu dalam Alfredo, Jondri dan Rismala, 2015).

Konsep SVR didasarkan pada *risk minimization*, yaitu mengestimasi suatu fungsi dengan cara meminimalkan batas atas dari *generalization error*, sehingga SVR mampu mengatasi *overfitting*. (Gunn, 1998).

Tugas dari regresi adalah untuk menemukan fungsi antara x_i dengan y_i (Zhao dkk, 2015) yang dalam kasus linier di tulis :

$$f(x) = w \cdot x + b \quad (2.39)$$

Pada kasus nonlinier, pemetaan nonlinier : $R^1 \rightarrow F$, dimana F merupakan ruang fitur dari ϕ yang diperkenalkan untuk menterjemahkan kerumitan masalah regresi nonlinier pada R^1 untuk sebuah masalah sederhana regresi linier pada F . Fungsi regresi setelah transformasi menjadi seperti berikut:

$$f(x) = w \cdot \Phi(x) + b \quad (2.40)$$

Dengan w merupakan vector pembobot, $\Phi(x)$ merupakan fungsi yang memetakan x dalam suatu dimensi dan b merupakan bias. Untuk mengevaluasi seberapa baik fungsi regresi, maka digunakan fungsi ε -insetive loss sebagai berikut.

$$L_\varepsilon(y, f(x)) = \begin{cases} 0 & \text{untuk } |y - f(x)| \leq \varepsilon \\ |y - f(x)| - \varepsilon & \text{dan sebaliknya} \end{cases} \quad (2.41)$$

Fungsi ε -insetive loss digunakan untuk mengukur risiko empiris, risiko empiris diukur berdasarkan persamaan (2.41), selisih target dengan hasil estimasi. Oleh karena itu parameter ε harus diatur untuk meminimalisir risiko empiris dengan menggunakan variabel slack ξ, ξ^* yang menggambarkan simpangan dari data *training* di luar zona ε -insetive.

Disamping meminimalisir kesalahan empiris dengan fungsi ε -insetive, juga harus meminimalisir norma Euclidean dari beban yang linier $\|w\|$ yang mana berhubungan dengan kemampuan generalisasi dari model SVR yang *ditraining* (Zhao dkk, 2015). Tujuannya adalah untuk memperlebar (*maximize*) *margin* sehingga kelandaian kurva serta kompleksitas model dapat dipastikan (Suga-

nyadevi dan Babulal, 2014). Sehingga permasalahan regresi dapat dinyatakan seperti masalah optimasi kuadratik berikut ini:

$$L(w, \xi) = \frac{1}{2} \|w\|^2 + C[\sum_{i=1}^n (\xi_{2i} + \xi_{2i}')] , \quad c > 0 \quad (2.42)$$

$$\text{subject to } \begin{cases} y_i - w * \Phi(x_i) - b \leq \varepsilon + \xi_i \\ w * \Phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \Phi_i, \xi_i^* \geq 0 \end{cases} \quad (2.43)$$

dimana C menyatakan koefisien penalti yang mendeterminasikan *trade-off* antara keempirisan dengan kesalahan generalisasi yang mana nilai C tersebut perlu diatur (Zhao dkk, 2015). Menurut Suganyadevi dan Babulal (2014) nilai parameter C yang optimal berada pada *range* 1-1000, sedangkan Huang dkk (2007) menyatakan nilai parameter C optimal pada *range* $10^{-2} - 10^4$. Untuk menyelesaikan permasalahan optimasi kuadratik pada persamaan (2.42), dapat digunakan dual Lagrangian:

$$f(x_i) = (w \Phi(x_i) + b) = \sum_{j=1}^n \alpha_j K(x_i, x_j) + b \quad (2.44)$$

dimana $K(x_i, x_j)$ merupakan fungsi kernel.

SVR dapat digunakan pada data non linier dengan menggunakan pendekatan kernel. Fungsi kernel digunakan untuk mengatasi kasus nonlinieritas pada data dengan cara mentransformasi data kedalam dimensi ruang yang lebih tinggi, sehingga dapat dipisahkan pada *feature space* yang baru. Pilihan fungsi kernel yang digunakan pada metode SVR sebagai berikut (Gunn, 1998).

1. Linier $x^T x$
2. Polinomial : $(x^T x + 1)^n$
3. *Radial Basis Function* (RBF) : $\exp(-\gamma \|x_i - x_j\|^2)$

Pada penelitian ini akan digunakan fungsi kernel *Radial Basis Function* (RBF). RBF merupakan fungsi kernel yang sering digunakan yang dapat mengatasi kasus nonlinieritas pada data dan menghasilkan hasil yang cukup bagus. Lippi, Bertini dan Frasconi (2013) dan Harafani (2015) sebelumnya telah membuktikan bahwa fungsi kernel RBF menghasilkan nilai akurasi peramalan yang lebih baik dibandingkan fungsi kernel lain. Jika pada pengujian nonlinieritas data menghasilkan kesimpulan bahwa data memiliki data pola linier, maka akan dilakukan penggantian fungsi kernel men-

jadi fungsi kernel linier. Fungsi kernel sangat penting dalam menentukan tingkat keakuratan dari prediksi. Jika fungsi kernel yang digunakan kernel *Radial Basis Function* (RBF) maka perlu mengatur parameter γ sebagai berikut.

$$K(x_i, x_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right) \quad (2.45)$$

Berdasarkan uraian mengenai SVR dapat dilihat bahwa terdapat tiga parameter bebas C , ε , γ . Nilai parameter tersebut harus diatur untuk mendapatkan hasil yang optimal yakni model yang menghasilkan error yang terkecil. Menurut Huang dkk (2007) *range* parameter γ berada pada *range* 0,002-3,8. Suganyadevi dan Babulal (2014) menyatakan bahwa nilai parameter ε yang optimal adalah 0,0001 dan untuk parameter γ berada pada *range* 0,05-1. Jia dkk (2011) berpendapat bahwa *range* parameter yang optimal untuk C adalah 100-1000, untuk ε adalah 0,001-0,01 dan untuk γ adalah 0,5-1,5. Pemilihan *range* parameter sangat penting untuk memperoleh hasil yang optimal.

2.4 Genetic Algorithm (GA)

Genetic Algorithm (GA) pertama kali dikembangkan oleh John Holland dari Universitas Michigan pada tahun 1975. GA dikembangkan berdasarkan teori evolusi dengan prinsip seleksi alam. GA merepresentasikan sebuah teknik identifikasi pendekatan solusi untuk masalah optimasi. GA adalah sebuah algoritma optimasi metaheuristik yang berdasarkan populasi solusi dan menggunakan mekanisme spesifik yang terinspirasi dari evolusi secara biologi seperti individu (kromosom), reproduksi, mutasi, rekombinasi, seleksi, ketahanan *fitness* (Gorunescu, 2010 diacu dalam Putri, 2016).

Berdasarkan teori evolusi spesies yang menyertakan proses seleksi alam seperti reproduksi, persilangan, mutasi dan lainnya, GA memanipulasi populasi struktur simbolis, yang mewakili solusi, agar mendapatkan adaptasi yang terbaik yang menghasilkan solusi yang terbaik untuk suatu permasalahan. Sebuah solusi yang dibangkitkan dalam GA disebut kromosom, sedangkan kumpulan

dari kromosom disebut sebagai populasi. (Thomas, 2007 diacu dalam Petrus dkk, 2009).

Pengkodean adalah suatu teknik untuk menyatakan populasi awal sebagai kandidat solusi suatu masalah ke dalam suatu kromosom. Gen dan Cheng (2000) juga menjelaskan bahwa pengkodean merupakan kunci pokok persoalan, dalam melakukan pengkodean harus diperhatikan apakah dapat membangun pencarian genetik yang efektif (Yuliani, Irhamah dan Prastyo, 2014).

Fitness individu dalam *Genetic Algorithm* adalah nilai fungsi objektif untuk fenotipe. Untuk menghitung *fitness*, kromosom harus terlebih dahulu didekode dan fungsi tujuan harus dievaluasi. *Fitness* tidak hanya menunjukkan bagaimana solusi yang baik, tetapi juga berhubungan dengan seberapa dekat kromosom pada solusi optimum (Sivanandam dan Deepa, 2008 diacu dalam Yuliani dkk, 2014).

Menurut Gen dan Cheng (2000), metode seleksi yang sering digunakan adalah *roulette wheel*. Ide dasar seleksi adalah untuk menentukan probabilitas seleksi atau probabilitas kelangsungan hidup pada setiap kromosom proporsional dengan nilai *fitness*nya. Anggota populasi yang memiliki *fitness* tinggi akan bertahan hidup dan dapat bereproduksi, anggota populasi yang memiliki *fitness* rendah akan mati. Setelah dilakukan seleksi maka selanjutnya yaitu mengoperasikan kromosom dengan *Crossover* dan mutasi.

Crossover adalah operator *Genetic Algorithm* yang utama karena beroperasi pada dua kromosom pada suatu waktu dan membentuk *offspring* dengan mengkombinasikan dua bentuk kromosom. Cara sederhana untuk memperoleh *crossover* adalah dengan memilih suatu titik yang dipisahkan secara random dan kemudian membentuk *offspring* dengan cara mengkombinasikan segmen dari satu induk ke sebelah kiri dari titik yang dipisahkan dengan segmen dari induk yang lain ke sebelah kanan dari titik yang dipisahkan. Schaffer dkk (1989) menyatakan bahwa probabilitas *crossover* (P_c) yang optimal adalah pada *range* 0,75-0,95.

Keanekaragaman individu dalam populasi telah dihasilkan dengan menggunakan proses seleksi dan pindah silang. Dengan

kedua operator genetik tersebut dapat terjadi hilangnya struktur gen tertentu sehingga tidak bisa diperoleh kembali informasi yang terkandung didalamnya. Operator mutasi diperkenalkan sebagai cara untuk mengembalikan informasi yang hilang tersebut. Melalui mutasi, individu baru dapat diciptakan dengan melakukan pengubahan terhadap satu atau lebih nilai gen pada individu yang sama. Schaffer dkk (1989) menjelaskan bahwa probabilitas mutasi yang baik ketika berada pada *range* 0,05-0,01, sedangkan Waghoo dan Dhar (2013) menyatakan probabilitas mutasi yang optimal pada *range* 0-0,3. Grefenstette (1986) menyatakan bahwa probabilitas mutasi tidak boleh lebih dari 0,05, karena akan sulit mendapatkan hasil parameter optimal dari GA.

Pembentukan populasi baru dengan *crossover* dan mutasi memungkinkan adanya kromosom yang paling baik hilang. Oleh karena itu, untuk menjaga agar kromosom bernilai *fitness* terbaik tidak hilang selama evolusi, maka perlu dibuat satu atau beberapa kopinya (*elitism*) (Yuliani dkk, 2014).

2.5 Kriteria Pemilihan Model Terbaik

Pemilihan model terbaik melalui pendekatan *in sample* dan *out sample* dengan menggunakan RMSE (*Root Mean Square*) dan pendekatan *out sample* menggunakan SMAPE (*Symmetric Mean Absolute Percentage Error*) . Model dikatakan baik, jika model tersebut memiliki nilai RMSE dan SMAPE yang kecil. Berikut rumus RMSE dan SMAPE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} = \sqrt{MSE} \quad (2.46)$$

$$SMAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{(Y_t - \hat{Y}_t)}{(Y_t + \hat{Y}_t)/2} \right| \times 100 \quad (2.47)$$

dimana,

- n : banyaknya data.
- Y_t : Data *actual* pada waktu ke- t .
- \hat{Y}_t : Data ramalan pada waktu ke- t .

2.6 Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Satuan yang digunakan dalam mengukur curah hujan adalah millimeter (mm). Curah hujan 1 (satu) milimeter artinya adalah dalam luasan satu meter persegi tempat yang datar, tertampung air setinggi satu milimeter atau tertampung air sebanyak satu liter. Intensitas hujan merupakan banyaknya curah hujan persatuan jangka waktu tertentu. Unsur-unsur hujan yang harus diperhatikan dalam mempelajari curah hujan adalah jumlah curah hujan dan intensitas atau kekuatan tetesan hujan. Berdasarkan intensitas curah hujan dibedakan menjadi 3 yaitu hujan sedang yang berada diantara 20 dan 50 mm perhari, hujan lebat berada diantara 50 dan 100 mm perhari, dan hujan sangat lebat berada diatas 100 mm perhari (Suriadikusumah, 2007).

2.7 Penentuan Pola Tanam Padi dan Kalender Tanam Padi

Oldeman (1980) membuat sistem klasifikasi iklim yang dihubungkan dengan pertanian menggunakan unsur curah hujan. Klasifikasi ini telah diterapkan pada berbagai penelitian dan menunjukkan hasil yang bermanfaat dalam bidang pertanian. Kriteria dalam klasifikasi iklim didasarkan pada perhitungan bulan basah (BB), bulan lembab (BL), dan bulan kering (BK). Suatu bulan disebut sebagai bulan basah apabila mempunyai curah hujan bulanan lebih besar dari 200 mm, disebut bulan lembab apabila mempunyai curah hujan bulanan antara 100-200 mm, dan disebut bulan kering apabila curah hujan bulanan di bawah 100 mm. Batasan yang digunakan adalah kebutuhan air tanaman dan hujan efektif sebagai berikut.

1. Padi sawah membutuhkan air rata-rata per bulan 145 mm dalam musim hujan.
2. Palawija membutuhkan air rata-rata per bulan 50 mm dalam musim kemarau.
3. Hujan efektif untuk sawah adalah 100% kebutuhan air rata-rata per bulan.

Gambar 2.1 menunjukkan akan terjadi hujan pada bulan Januari hingga Mei. Pada bulan tersebut terdapat beberapa dasarian yang menunjukkan akan terjadi hujan lebat sehingga petani diharapkan mewaspadai hal tersebut. Musim tanam pertama dapat dilakukan hingga bulan April sedangkan musim tanam kedua pada bulan November, namun harus dibantu oleh sistem pengairan dikarenakan pada bulan November akan terjadi hujan sedang yang diduga tidak mencukupi kebutuhan air tanaman padi. Pada bulan Juni hingga Oktober, petani dapat mengganti padi dengan palawija karena pada bulan tersebut diramalkan hanya akan turun hujan dengan intensitas ringan sehingga palawija cocok ditanam karena membutuhkan air yang lebih sedikit.

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data curah hujan dasarian yang berasal dari Pos pengamatan curah hujan di Kedungadem dan Cawak Kabupaten Bojonegoro yang diambil di BMKG Karangploso Malang. Periode data adalah antara bulan Januari 2000-Juni 2016. Data bulan Januari 2000 hingga Desember 2015 digunakan sebagai data *in sample* dan data Januari 2016- Juni 2016 digunakan sebagai data *out sample*.

3.2 Variabel Penelitian

Berikut ini adalah variabel bebas dan variabel tak bebas yang digunakan dalam penelitian ini.

1. Variabel tak bebas (Y_t) adalah data curah hujan dasarian yang dijadikan data *actual/target*.
2. Variabel bebas (X) adalah data curah hujan dasarian pada waktu sebelumnya ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$).

Tabel 3. 1 Struktur Data

Tahun	Bulan	Dasarian	Curah Hujan	Keterangan
2000	Januari	1	Y_1	<i>In sample</i>
		2	Y_2	<i>In sample</i>
		3	Y_3	<i>In sample</i>
	:	:	:	:
	Desember	3	Y_{36}	<i>In sample</i>
:	:	:	:	:
2015	Desember	3	Y_{576}	<i>In sample</i>
2016	Januari	1	Y_{577}	<i>out sample</i>
	:			
	Juni	3	Y_{594}	<i>out sample</i>

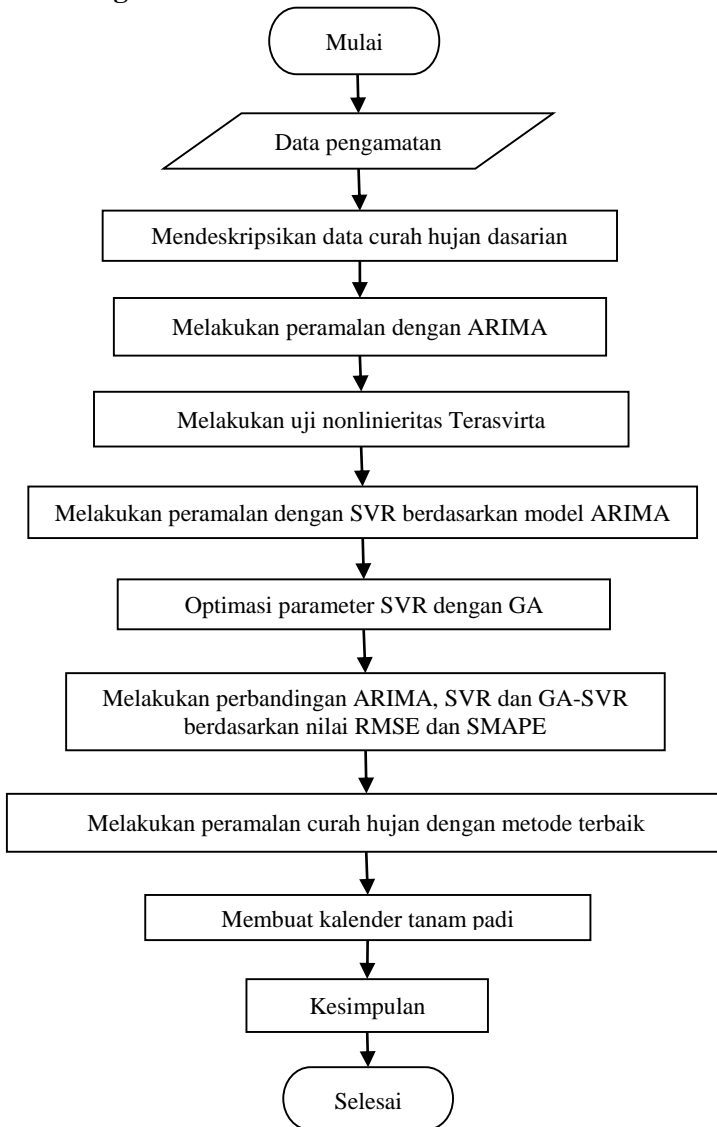
3.3 Langkah Analisis

1. Melakukan analisis statistika deskriptif dan plot data curah hujan dasarian di Pos Kedungadem dan Cawak.
2. Analisis ARIMA
 Peramalan curah hujan dengan metode ARIMA dilakukan dengan prosedur sebagai berikut.
 - a. Mengidentifikasi kestasioneran data terhadap *mean* dan varians menggunakan plot *Box-Cox* dan pengujian ADF. Jika data menunjukkan belum stasioner dalam varians, maka dilakukan transformasi. Jika data belum stasioner dalam *means* maka dilakukan *differencing*.
 - b. Mengidentifikasi model ARIMA sementara berdasarkan plot ACF dan PACF.
 - c. Melakukan estimasi parameter dan pengujian signifikansi parameter. Jika parameter belum signifikan maka kembali ke langkah 2b.
 - d. Melakukan *diagnostic checking* dari model yang sudah memiliki parameter yang signifikan. Beberapa asumsi residual yang harus dipenuhi adalah asumsi *white noise* dan asumsi distribusi normal. Jika terdapat asumsi yang belum terpenuhi maka kembali ke langkah 2b.
 - e. Mengukur kebaikan model dalam melakukan peramalan berdasarkan RMSE dan SMAPE.
3. Uji non linieritas Terasvirta
 - a. Meregresikan Y_t dengan $1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$ dan menghitung residual \hat{a}_t serta menghitung jumlah kuadrat residual.
 - b. Meregresikan \hat{a}_t dengan $1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$ dan m prediktor tambahan yang merupakan suku kubik atau kuadratik yang merupakan hasil pendekatan ekspansi deret Taylor.
 - c. Menghitung nilai F dan membandingkannya dengan P -value.

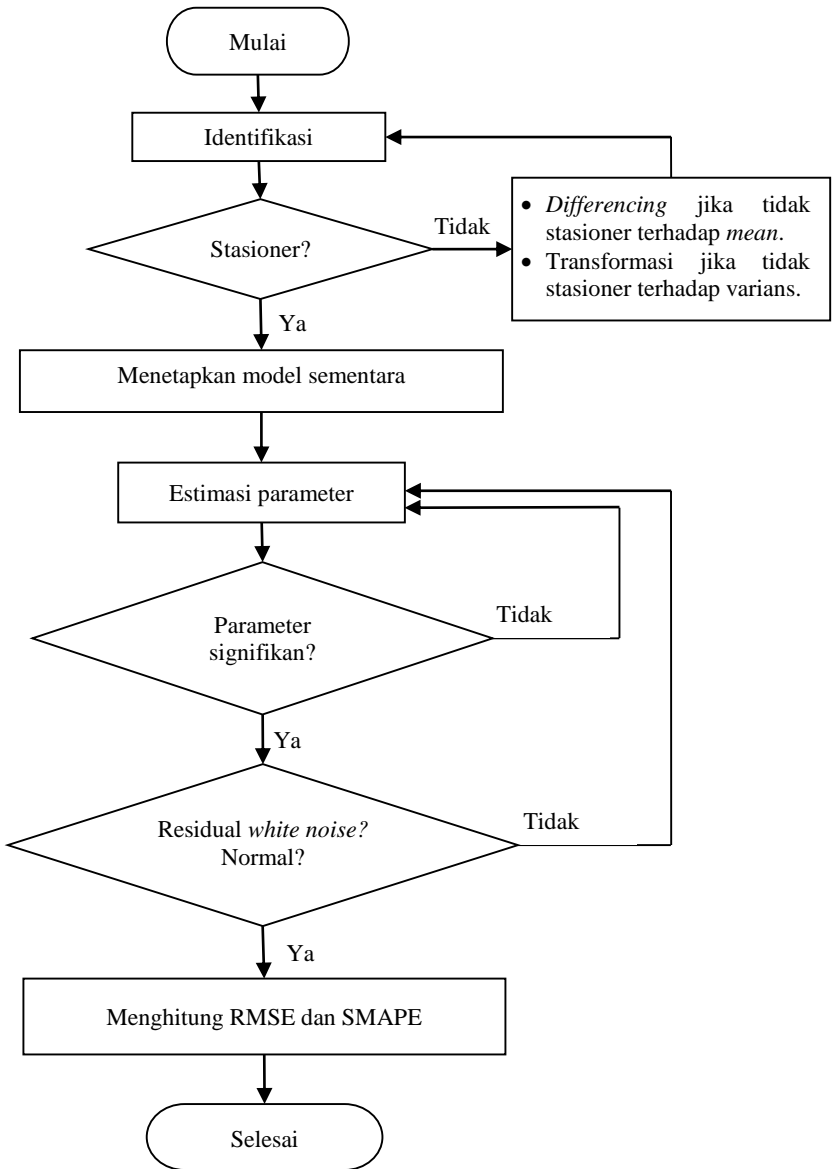
4. Analisis dengan metode *Support Vector Regression* (SVR).
 - a. Menyusun model SVR dengan variable bebas (X) adalah data curah hujan dasarian pada waktu sebelumnya yang diperoleh dari penguraian model ARIMA yang sudah memiliki parameter yang signifikan pada langkah 2
 - b. Membagi data *in sample* dan *out sample* sesuai dengan hasil model regresi pada langkah 4a.
 - c. Menentukan fungsi kernel yang akan digunakan. Pada penelitian ini menggunakan fungsi kernel *radial basis function* (RBF).
 - d. Menentukan *range* nilai parameter C , ε dan γ untuk optimasi *hyperplane* pada data *in sample*.
 - e. Melakukan pemodelan dengan SVR berdasarkan *range* nilai parameter.
 - f. Melakukan pengecekan terhadap solusi yang didapatkan. Model terbaik adalah model yang memiliki nilai parameter paling optimum yang memiliki nilai *error* terkecil, jika kondisi ini belum terpenuhi maka proses diulang dari langkah 4e.
 - g. Melakukan peramalan berdasarkan model SVR terbaik .
 - h. Mengukur kebaikan model dalam melakukan peramalan berdasarkan RMSE dan SMAPE.
5. Analisis dengan metode *Genetic Algorithm-Support Vector Regression* (GA-SVR).
 - a. Menyusun kromosom dengan membangkitkan 100 kromosom. Kromosom yang dibangkitkan terdiri dari tiga gen yang menunjukkan parameter model SVR dengan menggunakan fungsi kernel RBF. Nilai inisial kromosom berasal dari nilai parameter model SVR yang telah didapatkan pada langkah 4.
 - b. Menentukan nilai *fitness*. Pada penelitian ini nilai *fitness* yang digunakan adalah kesalahan dalam peramalan yaitu nilai RMSE.

- c. Melakukan proses seleksi sebanyak 100 kromosom dari sejumlah 100 induk yang berasal dari populasi dengan seleksi *roulette wheel*.
 - d. Melakukan proses pindah silang jika nilai bilangan *random* yang dibangkitkan kurang dari probabilitas pindah silang (P_c) = 0,8.
 - e. Melakukan proses mutasi jika nilai bilangan *random* yang dibangkitkan kurang dari nilai probabilitas mutasi (P_m) = 0,01.
 - f. Melakukan proses *elitism*.
 - g. Melakukan pergantian populasi lama dengan generasi baru dengan cara memilih sejumlah kromosom yang memiliki nilai *fitness* terbaik yang telah melalui proses seleksi, pindah silang dan *elitism*.
 - h. Melakukan pengecekan terhadap solusi yang telah didapatkan. Solusi dikatakan telah mencapai kriteria apabila nilai *fitness* terbaik telah konvergen jika kondisi ini belum terpenuhi maka proses *Genetic Algorithm* diulang dari langkah 5c.
 - i. Melakukan estimasi parameter dengan metode GA.
 - j. Memasukkan parameter yang didapat dari GA ke dalam algoritma SVR.
 - k. Mengukur kebaikan model dalam melakukan peramalan berdasarkan RMSE dan SMAPE.
6. Melakukan perbandingan metode ARIMA, SVR dan GA-SVR berdasarkan nilai RMSE dan SMAPE untuk memilih metode peramalan terbaik
 7. Melakukan peramalan curah hujan 6 bulan selanjutnya dengan metode terbaik.
 8. Menyusun kalender tanam padi berdasarkan pada hasil peramalan 6 bulan selanjutnya.
 9. Membuat kesimpulan.

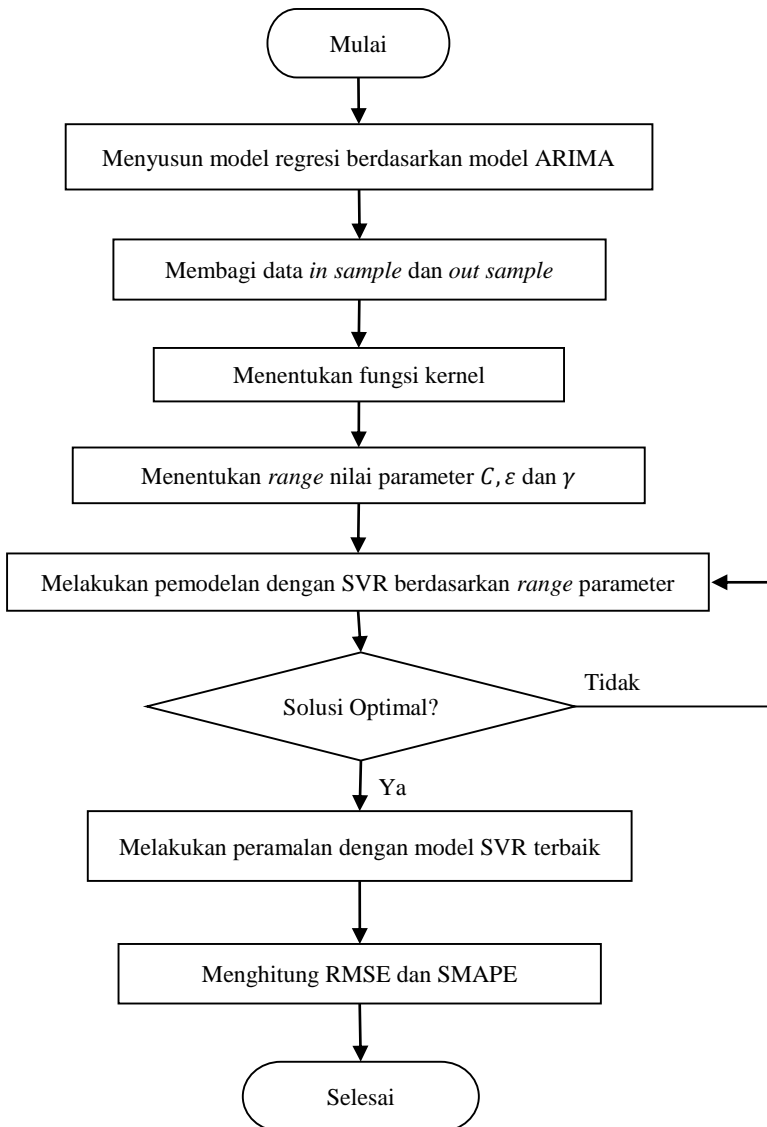
3.4 Diagram Alir



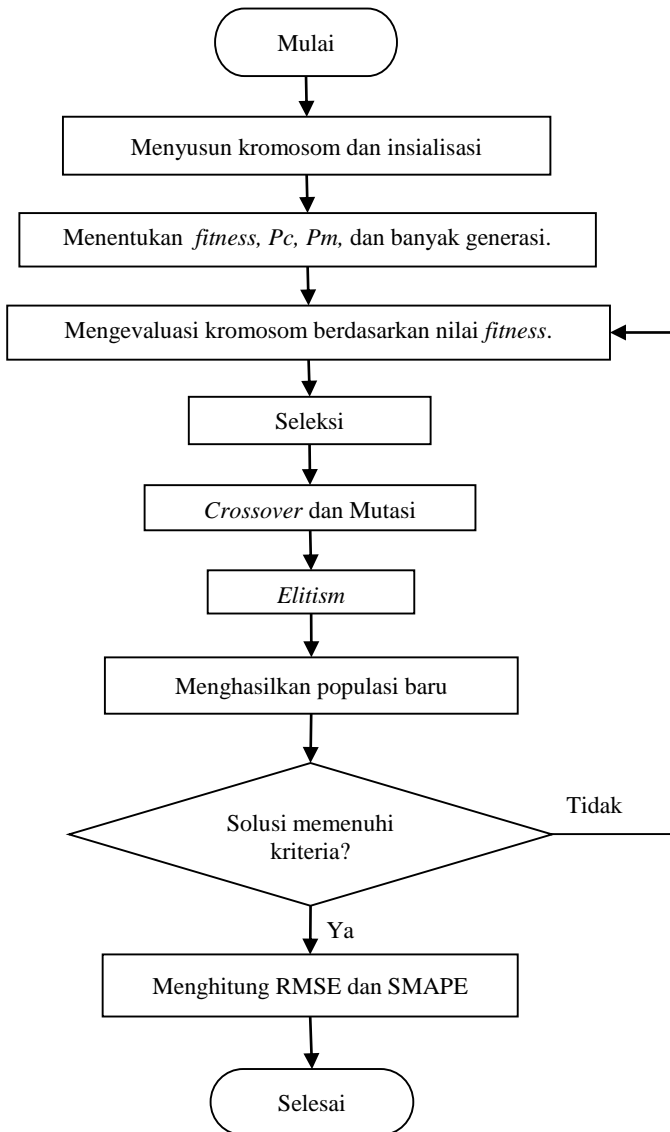
Gambar 3. 1 Diagram Alir Penelitian



Gambar 3. 2 Diagram Alir Peramalan Menggunakan ARIMA



Gambar 3. 3 Diagram Alir Peramalan Menggunakan SVR



Gambar 3. 4 Diagram Alir Peramalan Menggunakan GA-SVR

BAB IV

ANALISIS DAN PEMBAHASAN

Pada bab ini akan dibahas dan dijelaskan mengenai prosedur pembentukan model peramalan curah hujan yang dilakukan di dua pos pengamatan di Kabupaten Bojonegoro, yaitu Pos Cawak dan Kedungadem dengan metode ARIMA, SVR dan GA-SVR. Setelah mendapatkan model peramalan dari ketiga metode tersebut, maka hasil dari metode tersebut akan dibandingkan untuk memperoleh model peramalan terbaik yang akan digunakan untuk meramalkan curah hujan di dua pos pengamatan tersebut.

4.1 Karakteristik Curah Hujan Kabupaten Bojonegoro

Deskripsi curah hujan penting dilakukan untuk mengetahui karakteristik-karakteristik khusus curah hujan pada pos pengamatan yaitu Pos Cawak dan Kedungadem. Data yang digunakan adalah data curah hujan dasarian mulai bulan Januari 2000 hingga bulan Juni 2016. Jenis eksplorasi data yang dilakukan yaitu menggunakan statistika deskriptif baik berupa numerik maupun grafis. Hasil statistika deskriptif curah hujan dasarian pada kedua pos pengamatan ditunjukkan pada Tabel 4.1.

Tabel 4. 1 Statistika Deskriptif Data Curah Hujan Dasarian

Pos	Mean	Varians	Maks	Skewness	Kurtosis
Cawak	40,85	2546,98	278	1,47	1,99
Kedungadem	52,33	3681,13	357	1,47	2,31

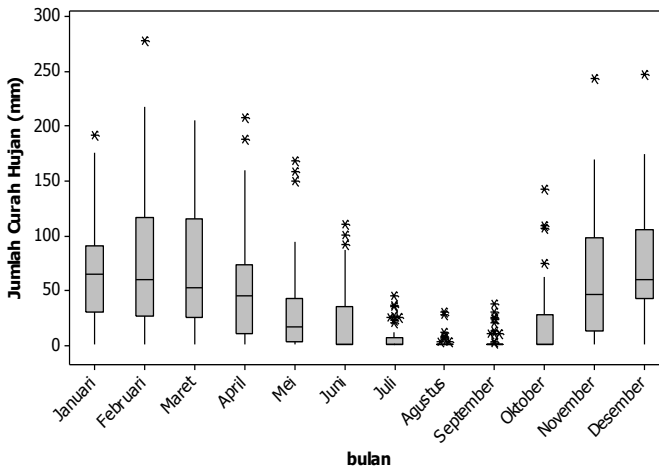
Tabel 4.1 menunjukkan bahwa rata-rata curah hujan dasarian pada kedua pos cukup berbeda. Rata-rata curah hujan pada Pos Kedungadem lebih besar yaitu 52,23 mm per dasarian, hal ini termasuk pada kategori curah hujan lebat. Sementara pada Pos Cawak rata-rata curah hujan termasuk pada kategori hujan sedang, karena jumlah curah hujan berada pada *range* 20-50 mm. Pada ukuran keragaman dapat dilihat dari nilai varians. Jika nilai varians semakin tinggi menunjukkan bahwa data tersebut semakin beragam. Keragaman data curah hujan tertinggi terdapat pada Pos Kedungadem yang menunjukkan bahwa curah hujan di daerah sekitar Pos Ke-

dungadem cenderung memiliki intensitas curah hujan yang lebih berfluktuatif dibandingkan pada daerah sekitar Pos Cawak.

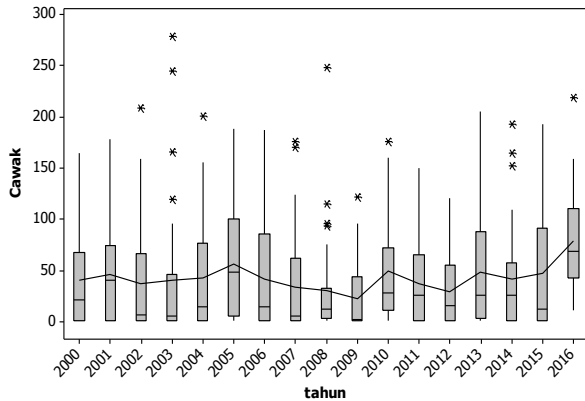
Persebaran data dapat dilihat dari *skewness* dan *kurtosis*. Nilai *skewness* dan *kurtosis* menunjukkan apakah suatu data memiliki bentuk kurva yang mengikuti kurva distribusi normal. Nilai *skewness* pada kedua pos pengamatan sama dan bernilai positif. Hal ini menunjukkan tingkat kemiringan grafik persebaran data condong ke kanan, sehingga dapat dinyatakan grafik persebaran data pada kedua pos pengamatan tidak mengikuti distribusi normal. Selain itu, pada nilai *kurtosis* menunjukkan nilai yang positif yang menunjukkan grafik persebaran data lebih runcing dibandingkan grafik persebaran data yang mengikuti distribusi normal.

4.1.1 Deskripsi Curah Hujan Di Pos Cawak

Untuk melihat kondisi curah hujan tiap bulan dan tiap tahun pada Pos Cawak dapat dilihat secara visual dengan menggunakan *boxplot* seperti pada Gambar 4.1 dan Gambar 4.2.



Gambar 4. 1 *Boxplot* Curah Hujan Per Bulan Pos Cawak



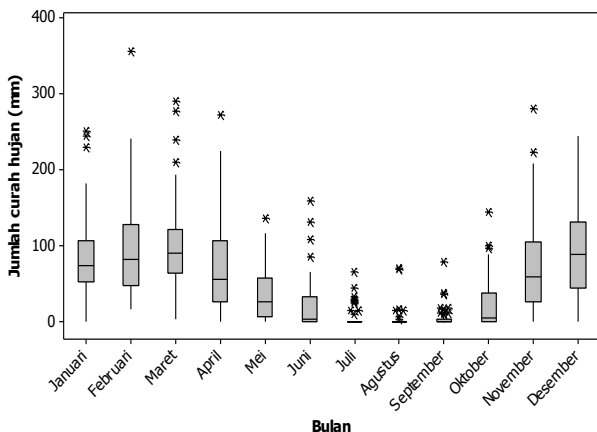
Gambar 4. 2 Boxplot Curah Hujan Per Tahun Pos Cawak

Gambar 4.1 menunjukkan bahwa pada bulan Januari, Februari, Maret, April, November dan Desember memiliki nilai rata-rata curah hujan yang cukup tinggi. Sementara itu, pada bulan Juli, Agustus dan September memiliki nilai rata-rata curah hujan yang rendah, yang mengindikasikan bahwa terjadi musim kemarau dan musim pancaroba yaitu pergantian dari musim hujan ke musim kemarau atau sebaliknya pada bulan tersebut. Pada bulan Februari, November dan Desember terdapat *outlier* yang tinggi yang menunjukkan terjadinya hujan dengan intensitas yang tinggi.

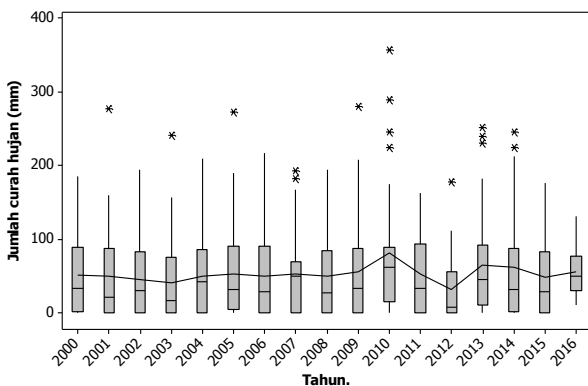
Berdasarkan Gambar 4.2 didapatkan informasi bahwa pada tahun 2016 memiliki nilai rata-rata curah hujan yang tertinggi selama 17 tahun terakhir. Sementara itu, pada tahun 2009 memiliki nilai rata-rata curah hujan yang terendah. Pola curah hujan dari tahun ke tahun tidak sama dan cenderung mengalami fluktuasi.

4.1.2 Deskripsi Curah Hujan Di Pos Kedungadem

Kondisi curah hujan tiap bulan dan tiap tahun pada Pos Kedungadem dapat dilihat secara visual dengan *boxplot* pada Gambar 4.3 dan Gambar 4.4.



Gambar 4. 3 Boxplot Curah Hujan Per Bulan Pos Kedungadem



Gambar 4. 4 Boxplot Curah Hujan Per Tahun Pos Kedungadem

Berdasarkan Gambar 4.3 menunjukkan bahwa rata-rata curah hujan yang cukup tinggi terjadi pada bulan Januari, Februari, Maret, April, November dan Desember. Pada bulan Mei, Juni, Juli, Agustus, September dan Oktober memiliki rata-rata curah hujan yang rendah, yang mengindikasikan bahwa pada bulan tersebut terjadi musim kemarau dan musim pancaroba yaitu pergantian dari musim hujan ke musim kemarau atau sebaliknya. Bulan Januari, Februari, Maret, April dan November terdapat *outlier* yang tinggi

yang menunjukkan terjadinya hujan dengan intensitas yang tinggi. Hal ini berbeda dengan Pos Cawak, dimana hujan dengan intensitas tinggi terjadi pada bulan Februari, November dan Desember. Jika dibandingkan dengan Pos Cawak, terdapat perbedaan waktu musim kemarau. Pada Pos Cawak musim kemarau terjadi pada bulan Agustus dan September, namun pada Pos Kedungadem diduga musim kemarau terjadi selama 3 bulan yaitu bulan Juli, Agustus dan Desember.

Gambar 4.4 menunjukkan bahwa selama 16 tahun terakhir rata-rata curah hujan tertinggi terjadi pada tahun 2010 dan tahun 2012 memiliki nilai rata-rata curah hujan yang terendah. Berdasarkan Gambar 4.4 menunjukkan pola hujan dari tahun ke tahun relatif sama, hal ini berbeda dengan Pos Cawak yang cenderung fluktuatif.

4.2 Peramalan Curah Hujan Pos Cawak

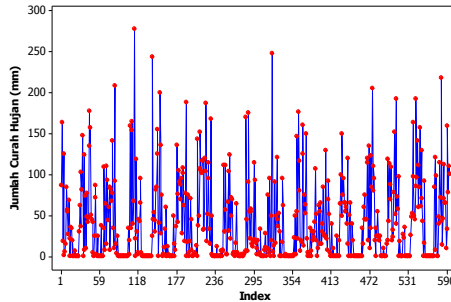
Pada tahap ini akan dilakukan peramalan curah hujan di Pos Cawak menggunakan tiga metode yaitu ARIMA, SVR dan GA-SVR. Peramalan dengan metode SVR dan GA-SVR akan didasarkan pada *input* dari model ARIMA yang sudah memiliki parameter yang signifikan. Setelah didapatkan model terbaik akan dilakukan pemilihan metode terbaik berdasarkan nilai RMSE dan SMAPE.

4.2.1 Peramalan Curah Hujan Menggunakan ARIMA

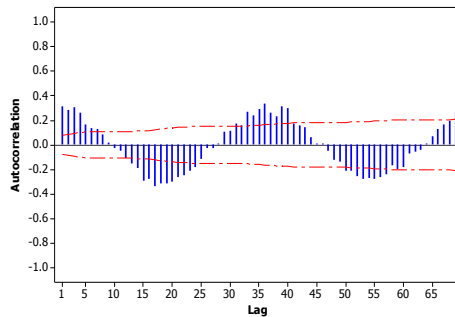
Pada peramalan dengan ARIMA terlebih dahulu akan dilakukan identifikasi terhadap pola data yang bertujuan untuk mengetahui apakah data curah hujan dasarian di Pos Cawak sudah memenuhi asumsi stasioner atau belum. Terdapat dua identifikasi kestasioneran data, yaitu stasioner dalam varians dan stasioner dalam *mean*. Dalam menentukan model peramalan curah hujan di Pos Cawak menggunakan metode ARIMA, perlu mengidentifikasi *time series plot* dari data tersebut untuk mengetahui pola dan karakteristik dari data

Gambar 4.5 menunjukkan bahwa curah hujan dasarian di Pos Cawak belum stasioner terhadap *mean* dan varians karena fluk-

tuasi antara titik satu dengan titik lainnya sangat bervariasi. Berdasarkan Plot ACF pada Gambar 4.6 didapatkan informasi bahwa data memiliki pola musiman (*seasonal*).



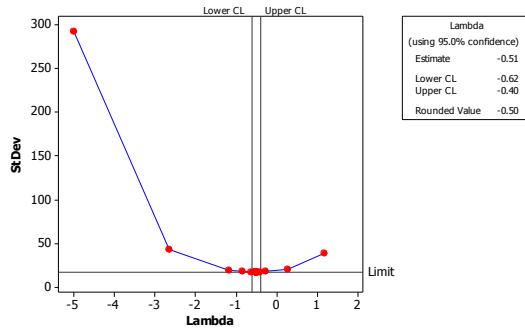
Gambar 4. 5 *Time Series Plot* Curah Hujan Dasarian Pos Cawak



Gambar 4. 6 Plot ACF Curah Hujan Dasarian Pos Cawak

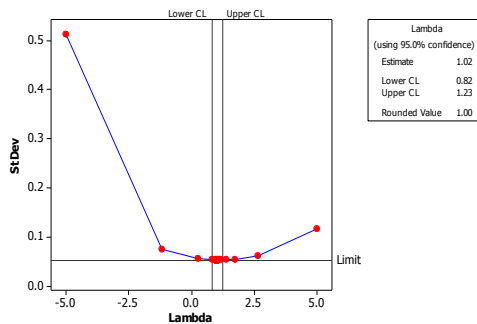
Setelah mengidentifikasi pola dan karakteristik data, langkah selanjutnya dilakukan analisis dengan melihat plot *Box-Cox* untuk melihat apakah data curah hujan dasarian sudah stasioner terhadap varians seperti pada Gambar 4.7

Berdasarkan Gambar 4.7 diketahui bahwa data curah hujan dasarian di Pos Cawak belum stasioner terhadap varians. Hal ini ditunjukkan dengan nilai *Lower CL* dan *Upper CL* yang belum melewati angka 1 atau bisa juga dilihat pada nilai *rounded value* (λ) yang belum bernilai satu. Karena nilai λ sebesar -0,5 maka data akan ditransformasikan $1/\sqrt{Y_t}$ untuk menstabilkan varians.



Gambar 4. 7 *Box-Cox* Data Curah Hujan Dasarian Pos Cawak

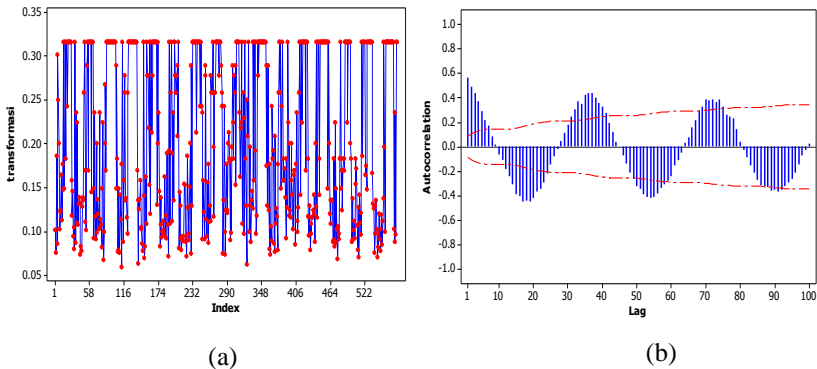
Setelah dilakukan transformasi maka dilakukan analisis *Box-Cox* kembali untuk mengetahui apakah data sudah stasioner terhadap varians atau belum. Gambar 4.8 merupakan *Box-Cox* curah hujan dasarian yang telah dilakukan transformasi. Dapat dilihat bahwa nilai *Lower CL* dan *Upper CL* yang sudah melewati angka 1 dan nilai *rounded value* (λ) yang bernilai satu, sehingga data hasil transformasi sudah stasioner terhadap varians.



Gambar 4. 8 *Box-Cox* Data Curah Hujan Dasarian Pos Cawak Hasil Transformasi

Setelah stasioneritas data terhadap varians terpenuhi, maka selanjutnya dilakukan pengujian stasioneritas data terhadap *mean* dengan menggunakan uji ADF. Uji ADF dilakukan dalam be-

berapa *lag* yakni *lag* 1, *lag* 2, *lag* 5, *lag* 8, *lag* 15 dan *lag* 30, dimana dari beberapa *lag* tersebut didapatkan hasil *p-value* yang sama yakni sebesar 0,01. Nilai ini menunjukkan $P\text{-value} < \alpha$, sehingga tolak H_0 . Dapat disimpulkan bahwa data transformasi sudah stasioner terhadap *mean*. Berikut adalah *time series plot* dan plot ACF data transformasi.



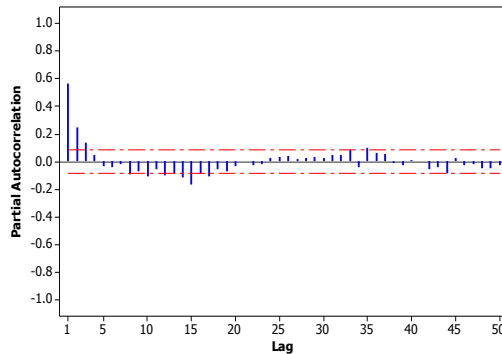
Gambar 4. 9 Time Series Plot (a) dan Plot ACF (b) Data Hasil Transformasi Pos Cawak

Berdasarkan Gambar 4.9 (a) diperoleh informasi bahwa data transformasi secara visual belum stasioner terhadap *mean* yang ditunjukkan plot yang tidak berfluktuasi disekitar garis *mean*. Pada plot ACF data transformasi menunjukkan bahwa data memiliki pola musiman. Jika berdasarkan pada uji ADF maka data hasil transformasi telah stasioner terhadap *mean*, namun hasil ini berbeda dengan *time series plot*. Oleh karena itu, untuk analisis ARIMA pada data transformasi akan dibagi menjadi ARIMA tanpa *differencing*, ARIMA *differencing* musiman 18 dan ARIMA *differencing* musiman 36.

A. ARIMA Pos Cawak Tanpa *Differencing*

Setelah data hasil transformasi sudah stasioner terhadap varians dan *mean* maka selanjutnya dilakukan identifikasi model ARIMA sementara berdasarkan pada plot ACF dan plot PACF data hasil transformasi sebagai berikut.

Pada plot ACF (Gambar 4.9 (b)) menunjukkan bahwa plot ACF memiliki pola *dies down*, sedangkan plot PACF pada Gambar 4.10 diperoleh informasi terdapat beberapa *lag* yang keluar dari batas *significance limit* yang menunjukkan bahwa *lag* tersebut signifikan. *Lag* tersebut adalah *lag 1, lag 2, lag 3, lag 8, lag 10, lag 12, lag 14, lag 15, lag 17, lag 33 dan lag 35*, sehingga terdapat beberapa kemungkinan model ARIMA yang didapatkan. Beberapa model dugaan ARIMA berdasarkan plot ACF dan PACF adalah ARIMA(3,0,0), ARIMA([1,2,3,15],0,0), ARIMA([8,10],0,0), ARIMA ([10,12,35],0,0) dan ARIMA ([1,2,3,35],0,0).



Gambar 4. 10 Plot PACF Data Transformasi Curah Hujan Pos Cawak

Setelah menetapkan beberapa model dugaan ARIMA, maka selanjutnya adalah mengestimasi parameter-parameter model dengan menggunakan *Conditional Least Square* (CLS). Kemudian parameter-parameter tersebut akan di uji signifikansinya menggunakan uji *t*. Hasil estimasi parameter dan pengujian signifikansi parameter model dugaan ARIMA di Pos Cawak tanpa melakukan *differencing* terdapat pada Tabel 4.2.

Berdasarkan Tabel 4.2 didapatkan hasil bahwa terdapat satu model yang memiliki parameter yang tidak signifikan yang ditunjukkan oleh nilai *P-value* yang lebih dari α ($\alpha=0,05$). Dari 5 model dugaan ARIMA terdapat 4 model yang memiliki parameter yang sudah signifikan yaitu ARIMA (3,0,0), ARIMA ([8,10],0,0), ARIMA ([10,12,35],0,0) dan ARIMA ([1,2,3,35],0,0).

Tabel 4. 2 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak Tanpa *Differencing*

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (3,0,0)	$\hat{\phi}_1 = 0,4821$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,2643$	<0,0001	Signifikan
	$\hat{\phi}_3 = 0,2212$	<0,0001	Signifikan
ARIMA ([1,2,3,15],0,0)	$\hat{\phi}_1 = 0,4802$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,2612$	<0,0001	Signifikan
	$\hat{\phi}_3 = 0,2160$	<0,0001	Signifikan
	$\hat{\phi}_{15} = 0,0125$	0,6075	Tidak Signifikan
ARIMA ([8,10],0,0)	$\hat{\phi}_8 = 0,5845$	<0,0001	Signifikan
	$\hat{\phi}_{10} = 0,2889$	<0,0001	Signifikan
ARIMA ([10,12,35],0,0)	$\hat{\phi}_{10} = 0,2168$	<0,0001	Signifikan
	$\hat{\phi}_{12} = 0,1308$	0,0040	Signifikan
	$\hat{\phi}_{12} = 0,6367$	<0,0001	Signifikan
ARIMA ([1,2,3,35],0,0)	$\hat{\phi}_1 = 0,4137$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,1890$	<0,0001	Signifikan
	$\hat{\phi}_3 = 0,1698$	<0,0001	Signifikan
	$\hat{\phi}_{35} = 0,2200$	<0,0001	Signifikan

Selanjutnya akan dilakukan *diagnostic checking* untuk mengetahui apakah residual sudah memenuhi asumsi *white noise* dan berdistribusi normal. Pengujian asumsi residual yang *white noise* menggunakan Uji *Ljung-Box* sedangkan pengujian normalitas residual digunakan uji *Kolmogorov-smirnov* seperti pada Tabel 4.3 berikut.

Tabel 4. 3 *Diagnostic Checking* Model ARIMA Pos Cawak Tanpa *Differencing*

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (3,0,0)	6	0,0667	<0,0100
	12	0,4258	
	18	0,1749	
	24	0,1931	
	30	0,4818	
	36	0,0376	
	42	0,0054	
	48	0,0143	

Tabel 4. 3 *Diagnostic Checking Model ARIMA Pos Cawak Tanpa Differencing (Lanjutan)*

Model ARIMA	Uji <i>White Noise</i>		Uji Normalitas
	Lag	P-value	
ARIMA ([8,10],0,0)	6	<0,0001	<0,0100
	12	<0,0001	
	18	<0,0001	
	24	<0,0001	
	30	<0,0001	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	
ARIMA ([10,12,35],0,0)	6	<0,0001	<0,0100
	12	<0,0001	
	18	<0,0001	
	24	<0,0001	
	30	<0,0001	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	
ARIMA ([1,2,3,35],0,0)	6	0,0509	<0,0100
	12	0,1676	
	18	0,2438	
	24	0,4947	
	30	0,7849	
	36	0,2112	
	42	0,2053	
	48	0,2479	

Berdasarkan pengujian asumsi residual yang *white noise* pada Tabel 4.3, didapatkan informasi terdapat satu model dugaan ARIMA yang dapat memenuhi asumsi *white noise* yaitu ARIMA ([1,2,3,35],0,0). Hal ini ditunjukkan oleh nilai *P-value* yang dari lebih dari α ($\alpha=0,05$), sehingga gagal tolak H_0 . Pada Tabel 4.3 pada kolom uji normalitas telah didapatkan *P-value* dari uji *Kolmogorovsmirnov* pada residual. Dari pengujian normalitas residual didapatkan hasil model ARIMA yang sudah memenuhi asumsi *white noise*, tidak dapat memenuhi asumsi normalitas residual yang di-

tunjukkan oleh nilai P -value yang kurang dari α ($\alpha=0,05$), sehingga tolak H_0 . Dikarenakan model ARIMA tidak dapat memenuhi asumsi normalitas residual, maka akan dilakukan pendeteksian *outlier*. Setelah dilakukan pendeteksian *outlier* didapatkan sebanyak 10 *outlier* sebagai berikut

Tabel 4. 4 Deteksi *Outlier* Model ARIMA Pos Cawak Tanpa *Differencing*

Model ARIMA	Pengamatan ke-	Jenis <i>Outlier</i>
ARIMA ([1,2,3,35],0,0)	34	Additive
	481	Additive
	326	Additive
	149	Additive
	575	Shift
	61	Additive
	114	Additive
	440	Additive
	556	Additive
	161	Additive

Setelah mendapatkan pengamatan yang *outlier*, selanjutnya pengamatan yang *outlier* akan dimasukkan ke dalam model ARIMA sementara dan kemudian dilakukan estimasi parameter serta pengujian asumsi. Namun, hasil pemodelan dengan penambahan *outlier* tidak memberikan hasil residual yang berdistribusi normal, dimana p -value yang diperoleh sebesar $<0,0100$. Gejala ini tidak bisa diatasi dengan deteksi *outlier* dikarenakan histogram dari residual cenderung lebih runcing (leptokurtik) dan nilai kurtosisnya bernilai positif yang menunjukkan bahwa residual model ARIMA *outlier* tidak berdistribusi normal. Hasil pemodelan ARIMA dengan *outlier* terdapat pada Lampiran 5.

Setelah mendapatkan model ARIMA yang memiliki parameter signifikan dan memenuhi asumsi residual yang *white noise*, selanjutnya akan dilakukan perhitungan kriteria kebaikan model ARIMA berdasarkan pada RMSE dan SMAPE yang terdapat pada

Tabel 4.5. Model ARIMA ([1,2,3,35],0,0) dapat diuraikan sebagai berikut.

$$\begin{aligned}
 Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_{35} B^{35}) &= a_t \\
 Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_3 Y_{t-3}^* - \phi_{35} Y_{t-35}^* &= a_t \\
 Y_t^* &= a_t + \phi_1 Y_{t-1}^* + \phi_2 Y_{t-2}^* + \phi_3 Y_{t-3}^* + \phi_{35} Y_{t-35}^* \\
 Y_t^* &= 0,4137 Y_{t-1}^* + 0,1890 Y_{t-2}^* + 0,1698 Y_{t-3}^* + 0,2200 Y_{t-35}^* + a_t
 \end{aligned}$$

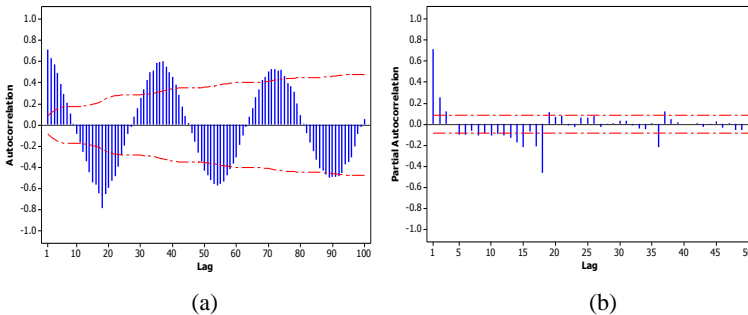
dimana Y_t^* adalah data curah hujan hasil transformasi

Tabel 4. 5 Kriteria Keباikan Model ARIMA Pos Cawak Tanpa *Differencing*

Model ARIMA	<i>In sample</i>	<i>Out sample</i>	
	RMSE	RMSE	SMAPE
ARIMA ([1,2,3,35],0,0)	51,4683	66,1683	75,3793

B. ARIMA Pos Cawak dengan *Differencing* Musiman 18

Time series plot data transformasi pada Gambar 4.9 (a) menunjukkan bahwa data hasil transformasi belum stasioner terhadap *mean*, serta plot ACF data tranformasi pada Gambar 4.9 (b) diketahui bahwa data memiliki pola musiman. Oleh karena itu akan dilakukan *differencing* musiman pada *lag* 18 terhadap data hasil transformasi. *Lag* tersebut dipilih karena diduga data curah hujan dasarian musiman selama 6 bulan. Gambar 4.11 menunjukkan plot ACF dan PACF data transformasi setelah dilakukan *differencing* musiman 18.



Gambar 4. 11 Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah *Differencing* Musiman 18 Pos Cawak

Dari plot ACF pada Gambar 4.11 (a) terlihat bahwa plot memiliki pola *dies down*. Sementara pada plot PACF (Gambar 4.11 (b)) terdapat beberapa *lag* yang keluar dari batas *significance limit* yaitu *lag* 1, *lag* 2, *lag* 3, *lag* 12, *lag* 13, *lag* 14, *lag* 15, *lag* 16, *lag* 17, *lag* 18 *lag* 19, *lag* 36 dan *lag* 37, sehingga terdapat lebih dari satu kemungkinan model yang didapatkan. Dugaan model ARIMA berdasarkan plot ACF dan PACF adalah ARIMA (3,0,0) (1,1,0)¹⁸, ARIMA (3,0,0)(2,1,0)¹⁸, ARIMA ([1,2,3,15],0,0)(2,1,0)¹⁸ dan ARIMA ([16,17,37],0,0)(1,1,0)¹⁸.

Setelah menetapkan beberapa model dugaan ARIMA, maka selanjutnya adalah mengestimasi parameter-parameter model dengan metode *conditional least square* (CLS). Kemudian parameter-parameter tersebut akan di uji signifikansinya menggunakan uji *t*. Berikut adalah estimasi parameter dan pengujian signifikansi parameter beberapa model dugaan ARIMA di Pos Cawak setelah *differencing* musiman 18.

Tabel 4. 6 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak dengan *Differencing* Musiman 18

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (3,0,0) (1,1,0) ¹⁸	$\hat{\phi}_1 = 0,2816$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,1885$	<0,0001	Signifikan
	$\hat{\phi}_3 = 0,1022$	0,0170	Signifikan
	$\hat{\Phi}_{18} = -0,6479$	<0,0001	Signifikan
ARIMA (3,0,0) (2,1,0) ¹⁸	$\hat{\phi}_1 = 0,3272$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,2076$	<0,0001	Signifikan
	$\hat{\phi}_3 = 0,1143$	0,0074	Signifikan
	$\hat{\Phi}_{18} = -0,7345$	<0,0001	Signifikan
	$\hat{\Phi}_{36} = -0,1937$	<0,0001	Signifikan
ARIMA ([1,2,3,15],0,0) (2,1,0) ¹⁸	$\hat{\phi}_1 = 0,3250$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,1905$	<0,0001	Signifikan
	$\hat{\phi}_3 = 0,1278$	0,0018	Signifikan
	$\hat{\phi}_{15} = -0,2400$	<0,0001	Signifikan
	$\hat{\Phi}_{18} = -0,6508$	<0,0001	Signifikan
	$\hat{\Phi}_{36} = -0,2478$	<0,0001	Signifikan

Tabel 4.6 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak dengan *Differencing* Musiman 18 (Lanjutan)

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA ([16,17,37],0,0) (1,1,0) ¹⁸	$\hat{\phi}_{16} = -0,0751$	0,1031	Tidak Signifikan
	$\hat{\phi}_{17} = -0,1268$	0,0073	Signifikan
	$\hat{\phi}_{37} = 0,0995$	0,0309	Signifikan
	$\hat{\Phi}_{18} = -0,7276$	<0,0001	Signifikan

Tabel 4.6 menunjukkan terdapat satu model yang memiliki parameter yang tidak signifikan yang ditunjukkan oleh nilai *P-value* yang lebih dari α ($\alpha=0,05$). Dari 4 model dugaan ARIMA terdapat 3 model yang sudah memiliki parameter yang signifikan yaitu ARIMA(3,0,0)(1,1,0)¹⁸, ARIMA(3,0,0)(2,1,0)¹⁸ dan ARIMA ([1,2,3,15],0,0) (1,1,0)¹⁸.

Setelah didapatkan 3 model dugaan ARIMA yang signifikan, maka analisis selanjutnya adalah *diagnostic checking* untuk mengetahui apakah residual sudah memenuhi asumsi *white noise* dan berdistribusi normal. Pengujian asumsi *white noise* menggunakan Uji *Ljung-Box*, sedangkan pengujian normalitas residual menggunakan uji *Kolmogorov-smirnov* seperti pada Tabel 4.7 berikut.

Tabel 4.7 *Diagnostic Checking* Model ARIMA Pos Cawak dengan *Differencing* Musiman 18

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (3,0,0)(1,1,0) ¹⁸	6	0,3494	>0,1500
	12	0,5990	
	18	0,0232	
	24	0,0031	
	30	0,0167	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	

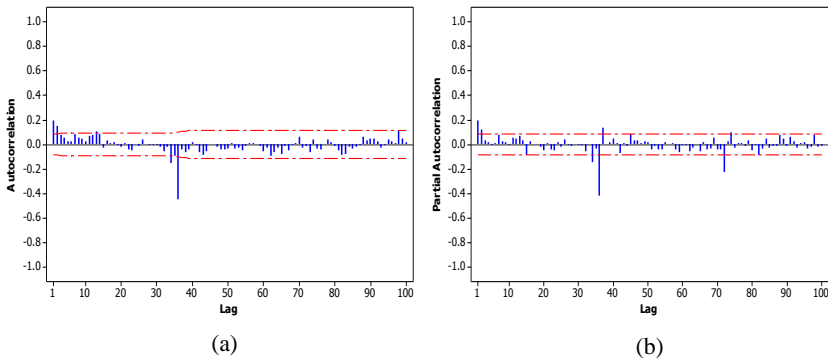
Tabel 4.7 *Diagnostic Checking Model ARIMA Pos Cawak dengan Differencing Musiman 18 (Lanjutan)*

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (3,0,0)(2,1,0) ¹⁸	6	0,0780	>0,1500
	12	0,4274	
	18	0,0045	
	24	0,0002	
	30	0,0012	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	
ARIMA ([1,2,3,15],0,0) (2,1,0) ¹⁸	12	0,1003	0,0122
	18	0,0417	
	24	0,0913	
	30	0,2264	
	36	0,0086	
	42	0,0039	
	48	0,0089	

Tabel 4.7 menunjukkan bahwa tidak terdapat model dugaan ARIMA yang dapat memenuhi asumsi *white noise* dan normalitas residual. Hal ini ditunjukkan oleh nilai *P-value* yang dari kurang dari α ($\alpha=0,05$), sehingga gagal tolak H_0 . Untuk analisis selanjutnya akan dilakukan *differencing* musiman 36 pada data hasil transformasi.

C. ARIMA Pos Cawak dengan *Differencing* Musiman 36

Gambar 4.9 (a) menunjukkan bahwa *time series plot* data hasil transformasi belum stasioner terhadap *mean*, sedangkan plot ACF data transformasi pada Gambar 4.9 (b) menunjukkan bahwa data memiliki pola musiman (*seasonal*). Oleh karena itu akan dilakukan *differencing* musiman pada lag 36 terhadap data hasil transformasi. Lag tersebut dipilih dikarenakan diduga data curah hujan dasarian musiman selama 12 bulan. Berikut adalah plot ACF dan PACF data hasil transformasi setelah dilakukan *differencing* musiman 36.



Gambar 4. 12 Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah *Differencing* Musiman 36 Pos Cawak

Gambar 4.12 (a) dan Gambar 4.12 (b) menunjukkan plot ACF dan PACF data hasil *differencing* musiman 36 dari data hasil transformasi. Dari plot ACF terlihat bahwa pada lag 1, lag 2, lag 13 dan lag 36 merupakan lag yang signifikan yang ditunjukkan oleh nilai ACF yang keluar dari batas *significance limit*. Sementara pada plot PACF terdapat beberapa lag yang keluar dari batas *significance limit* yaitu lag 1, lag 2, lag 34, lag 36, lag 37, lag 72 dan lag 74. Dugaan beberapa model ARIMA yaitu ARIMA (1,0,1) $(1,1,1)^{36}$, ARIMA(0,0,1)(0,1,1) 36 , ARIMA(0,0,2)(0,1,1) 36 dan ARIMA ([1,36,37],0,0) (0,1,0) 36 .

Langkah selanjutnya akan dilakukan estimasi parameter-parameter model dengan metode *Conditional Least Square* (CLS). Kemudian parameter-parameter tersebut di uji signifikansinya menggunakan uji *t*. Tabel 4.8 menunjukkan estimasi parameter dan pengujian signifikansi parameter model dugaan ARIMA di Pos Cawak setelah *differencing* musiman 36.

Tabel 4. 8 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak dengan *Differencing* Musiman 36

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (1,0,1) (0,1,1) 36	$\hat{\theta}_1 = 0,4511$	0,0051	Signifikan
	$\hat{\theta}_{36} = 0,7284$	<0,0001	Signifikan
	$\hat{\phi}_1 = 0,6264$	<0,0001	Signifikan

Tabel 4.8 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Cawak dengan *Differencing* Musiman 36 (Lanjutan)

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (0,0,1)(0,1,1) ³⁶	$\hat{\theta}_1 = -0,1690$	<0,0001	Signifikan
	$\hat{\theta}_{36} = 0,7313$	<0,0001	Signifikan
ARIMA (0,0,2)(0,1,1) ³⁶	$\hat{\theta}_1 = -0,1726$	<0,0001	Signifikan
	$\hat{\theta}_2 = -0,0888$	0,0415	Signifikan
	$\hat{\theta}_{36} = 0,7308$	<0,0001	Signifikan
ARIMA ([1,36,37],0,0) (0,1,0) ³⁶	$\hat{\phi}_1 = 0,2169$	<0,0001	Signifikan
	$\hat{\phi}_{36} = -0,4833$	<0,0001	Signifikan
	$\hat{\phi}_{37} = 0,14995$	0,0007	Signifikan

Berdasarkan hasil estimasi dan pengujian signifikansi parameter model dugaan ARIMA di Pos Cawak pada Tabel 4.8, didapatkan hasil bahwa semua model memiliki parameter yang signifikan yang ditunjukkan oleh nilai *P-value* yang kurang dari α ($\alpha=0,05$), sehingga tolak H_0 . Langkah selanjutnya adalah *diagnostic checking* untuk mengetahui apakah residual sudah memenuhi asumsi *white noise* dan berdistribusi normal. Pengujian asumsi *white noise* menggunakan Uji *Ljung-Box* sedangkan pengujian normalitas residual akan menggunakan uji *Kolmogorov-smirnov* seperti pada Tabel 4.9 berikut.

Tabel 4.9 *Diagnostic Checking* Model ARIMA Pos Cawak dengan *Differencing* Musiman 36

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (1,0,1)(0,1,1) ³⁶	6	0,7778	<0,0100
	12	0,7936	
	18	0,5961	
	24	0,8744	
	30	0,9752	
	36	0,9280	
	42	0,9756	
	48	0,9892	

Tabel 4.9 *Diagnostic Checking Model ARIMA Pos Cawak dengan Differencing Musiman 36 (Lanjutan)*

Model ARIMA	Uji <i>White Noise</i>		Uji Normalitas
	Lag	P-value	
ARIMA (0,0,1)(0,1,1) ³⁶	6	0,0700	<0,0100
	12	0,1016	
	18	0,0965	
	24	0,3196	
	30	0,6127	
	36	0,4999	
	42	0,6994	
	48	0,8030	
ARIMA (0,0,2)(0,1,1) ³⁶	6	0,2608	0,0229
	12	0,3711	
	18	0,2641	
	24	0,6022	
	30	0,8573	
	36	0,7601	
	42	0,8923	
	48	0,9412	
ARIMA ([1,36,37],0,0) (0,1,0) ³⁶	6	0,1391	<0,0100
	12	0,0717	
	18	0,0514	
	24	0,2134	
	30	0,5149	
	36	0,0517	
	42	0,0751	
	48	0,1285	

Tabel 4.9 menunjukkan bahwa semua model ARIMA sementara sudah memenuhi asumsi *white noise*, . Hal ini ditunjukkan oleh nilai *P-value* yang dari lebih dari α ($\alpha=0,05$), sehingga gagal tolak H_0 . Tabel 4.8 juga menunjukkan hasil *P-value* dari uji *Kolmogorov smirnov* pada residual. Dari pengujian normalitas residual didapatkan hasil dari model ARIMA yang memenuhi asumsi *white noise*, tidak terdapat model yang memenuhi asumsi normalitas residual yang ditunjukkan oleh nilai *P-value* yang kurang dari α ($\alpha=0,05$), sehingga tolak H_0 . Dikarenakan model ARIMA tidak

dapat memenuhi asumsi normalitas residual, maka analisis selanjutnya akan dilakukan pendeteksian *outlier*. Hasil pendeteksian pengamatan *outlier* terdapat pada Tabel 4.10.

Tabel 4. 10 Deteksi *Outlier* Model ARIMA Pos Cawak dengan *Differencing* Musiman 36

Model ARIMA	Pengamatan ke-	Jenis <i>Outlier</i>
ARIMA (1,0,1)(0,1,1) ³⁶	575	Shift
	334	Additive
	34	Additive
	114	Additive
	392	Additive
ARIMA (0,0,1)(0,1,1) ³⁶	575	Shift
	334	Additive
	34	Additive
	114	Additive
	209	Additive
ARIMA (0,0,2)(0,1,1) ³⁶	575	Shift
	334	Additive
	114	Additive
	440	Additive
	34	Additive
ARIMA ([1,36,37,0,0)(0,1,0) ³⁶	360	Shift
	575	Shift
	230	Shift
	114	Additive
	172	Additive

Tabel 4.10 menunjukkan hasil deteksi *outlier* pada 4 model dugaan ARIMA. Setelah mendapatkan pengamatan yang *outlier*, selanjutnya pengamatan *outlier* akan dimasukkan ke dalam model ARIMA sementara dan selanjutnya dilakukan estimasi parameter serta pengujian asumsi residual. Hasil pemodelan ARIMA dengan *outlier* terdapat pada Lampiran .

Hasil uji normalitas residual pada pemodelan ARIMA *outlier* pada Tabel 4.11 menunjukkan bahwa hasil pemodelan dengan penambahan *outlier* tidak memberikan hasil residual yang berdistribusi normal. Gejala ini tidak bisa diatasi dengan deteksi *outlier* dikarenakan histogram dari residual cenderung lebih runcing (leptokurtik) serta nilai kurtosisnya bernilai positif, hal ini menunjukkan residual model ARIMA dengan penambahan *outlier* tidak berdistribusi normal.

Tabel 4. 11 Hasi Uji Normalitas Residual Model ARIMA *Outlier* Pos Cawak dengan *Differencing* Musiman 36

No	Model	<i>P-value</i>
1	ARIMA (1,0,1)(0,1,1) ³⁶	<0,0100
2	ARIMA (0,0,1)(0,1,1) ³⁶	<0,0100
3	ARIMA (0,0,2)(0,1,1) ³⁶	<0,0100
4	ARIMA([1,36,37,0,0)(0,1,0) ³⁶	<0,0100

Selanjutnya akan dilakukan perhitungan kriteria kebaikan model ARIMA berdasarkan pada RMSE dan SMAPE untuk memilih model ARIMA terbaik. Perhitungan kriteria kebaikan model terdapat pada Tabel 4.12 sebagai berikut.

Tabel 4. 12 Kriteria Kebaikan Model ARIMA Pos Cawak dengan *Differencing* Musiman 36

No	Model ARIMA	<i>In sample</i>	<i>Out sample</i>	
		RMSE	RMSE	SMAPE
1	ARIMA (1,0,1)(0,1,1) ³⁶	48,4583	64,2174	73,9331
2	ARIMA (0,0,1)(0,1,1) ³⁶	49,9171	62,7889	68,3422
3	ARIMA (0,0,2)(0,1,1) ³⁶	49,4101	62,8525	70,4819
4	ARIMA ([1,36,37],0,0)(0,1,0) ³⁶	57,2861	66,0354	73,9663

Untuk menentukan model ARIMA terbaik dapat menggunakan kriteria *in sample* dan kriteria *out sample*. Berdasarkan Tabel 4.12 didapatkan hasil model terbaik adalah model ARIMA

(0,0,1) (0,1,1)³⁶ yang ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil dibandingkan model ARIMA lain. Berikut adalah penguraian model ARIMA yang terdapat pada Tabel 4.12

ARIMA (1,0,1)(0,1,1)³⁶

$$\begin{aligned} Y_t^* (1 - \phi_1 B)(1 - B^{36}) &= a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36} \\ Y_t^* - Y_{t-36}^* - \phi_1 Y_{t-1}^* + \phi_1 Y_{t-37}^* &= a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36} \\ Y_t^* &= Y_{t-36}^* + \phi_1 Y_{t-1}^* - \phi_1 Y_{t-37}^* + a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36} \\ Y_t^* &= Y_{t-36}^* + 0,6264 Y_{t-1}^* - 0,6264 Y_{t-37}^* + a_t - 0,4511 a_{t-1} \\ &\quad - 0,7284 a_{t-36} \end{aligned}$$

ARIMA (0,0,1)(0,1,1)³⁶

$$\begin{aligned} Y_t^* (1 - B^{36}) &= a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36} \\ Y_t^* - Y_{t-36}^* &= a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36} \\ Y_t^* &= Y_{t-36}^* + a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36} \\ Y_t^* &= Y_{t-36}^* + a_t + 0,1690 a_{t-1} - 0,7313 a_{t-36} \end{aligned}$$

ARIMA (0,0,2)(0,1,1)³⁶

$$\begin{aligned} Y_t^* (1 - B^{36}) &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{36} a_{t-36} \\ Y_t^* - Y_{t-36}^* &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{36} a_{t-36} \\ Y_t^* &= Y_{t-36}^* + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{36} a_{t-36} \\ Y_t^* &= Y_{t-36}^* + a_t + 0,1726 a_{t-1} + 0,0889 a_{t-2} - 0,7308 a_{t-36} \end{aligned}$$

ARIMA ([1,36,37],0,1)(0,1,1)³⁶

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_{36} B^{36} - \phi_{37} B^{37})(1 - B^{36}) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_{36} Y_{t-36}^* - \phi_{37} Y_{t-37}^* - Y_{t-36}^* + \phi_1 Y_{t-37}^* + \phi_{36} Y_{t-72}^* \\ &\quad + \phi_{37} Y_{t-73}^* = a_t \\ Y_t^* &= \phi_1 Y_{t-1}^* + \phi_{36} Y_{t-36}^* + \phi_{37} Y_{t-37}^* + Y_{t-36}^* - \phi_1 Y_{t-37}^* - \phi_{36} Y_{t-72}^* \\ &\quad - \phi_{37} Y_{t-73}^* + a_t \\ Y_t^* &= 0,2169 Y_{t-1}^* - 0,4833 Y_{t-36}^* + 0,14995 Y_{t-37}^* + Y_{t-36}^* \\ &\quad - 0,2169 Y_{t-37}^* + 0,4833 Y_{t-72}^* - 0,14995 Y_{t-73}^* + a_t \end{aligned}$$

dengan Y_t^* adalah data hasil transformasi

4.2.2 Peramalan Curah Hujan Menggunakan SVR

Sebelum masuk pada metode *Support Vector Regression* (SVR) terlebih dahulu akan dilakukan pengujian nonlinieritas pada data curah hujan di Pos Cawak. Pengujian nonlinieritas penting dilakukan untuk memilih fungsi kernel yang akan digunakan pada metode SVR. Uji nonlinieritas akan menggunakan uji Terasvirta. Dari uji Terasvirta didapatkan P -value sebesar $2,2 \cdot 10^{-6}$, sehingga

P -value kurang dari α ($\alpha=5\%$). Pada uji Terasvirta ini dilakukan dalam beberapa *lag* yaitu *lag* 1, *lag* 2, *lag* 3, *lag* 4, *lag* 5 dan *lag* 10, dari hasil pengujian tersebut didapatkan hasil P -value kurang dari α ($\alpha=5\%$), sehingga data curah hujan dasarian pada Pos Cawak memiliki pola non linier. Oleh karena itu dalam melakukan peramalan dengan SVR akan menggunakan fungsi kernel *Radial Basis Function* (RBF).

Pada peramalan dengan SVR akan menggunakan *lag-lag input* yang didapatkan dari penguraian model ARIMA yang sudah memiliki parameter yang signifikan. Berdasarkan pada penguraian model ARIMA pada analisis sebelumnya dan pada lampiran 19 diperoleh beberapa *lag* Y_t seperti pada Tabel 4.13.

Tabel 4. 13 *Input Lag SVR Pos Cawak*

No	Model	Input Lag (Y_t)
1	ARIMA (3,0,0)	$Y_{t-1}, Y_{t-2}, Y_{t-3}$
2	ARIMA ([8,10],0,0)	Y_{t-8}, Y_{t-10}
3	ARIMA([10,12,35],0,0)	$Y_{t-10}, Y_{t-12}, Y_{t-35}$
4	ARIMA([1,2,3,35],0,0)	$Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-35}$
5	ARIMA (3,0,0)(1,1,0) ¹⁸	$Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19},$ $Y_{t-2}, Y_{t-21}, Y_{t-37}, Y_{t-36},$ Y_{t-38}, Y_{t-39}
6	ARIMA (3,0,0) (2,1,0) ¹⁸	$Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19},$ $Y_{t-20}, Y_{t-21}, Y_{t-36}, Y_{t-37},$ $Y_{t-38}, Y_{t-39}, Y_{t-54}, Y_{t-55},$ Y_{t-56}, Y_{t-57}
7	ARIMA ([1,2,3,15],0,0)(2,1,0) ¹⁸	$Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-15}, Y_{t-18},$ $Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-33},$ $Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-39},$ $Y_{t-51}, Y_{t-54}, Y_{t-55}, Y_{t-56},$ Y_{t-57}, Y_{t-69}
8	ARIMA (1,0,1)(0,1,1) ³⁶	$Y_{t-1}, Y_{t-36}, Y_{t-37}$
9	ARIMA (0,0,1) (0,1,1) ³⁶	Y_{t-36}
10	ARIMA (0,0,[1,2]) (0,1,1) ³⁶	Y_{t-36}
11	ARIMA ([1,36,37],0,0) (0,1,0) ³⁶	$Y_{t-1}, Y_{t-36}, Y_{t-37}, Y_{t-72},$ Y_{t-73}

Berdasarkan pada Tabel 4.2, Tabel 4.6 dan Tabel 4.8 diperoleh 11 model ARIMA Pos Cawak yang telah memiliki parameter yang signifikan. Tabel 4.13 menunjukkan *input lag* yang didapatkan dari penguraian model ARIMA. Berdasarkan *input lag* tersebut, maka *lag* tersebut akan menjadi *input* (X) pada peramalan dengan SVR. Setelah mendapatkan *input*, maka tahap selanjutnya adalah menentukan parameter model SVR.

Pada metode SVR dengan menggunakan fungsi kernel RBF terdapat 3 parameter yaitu C (Cost), γ dan ε . Namun, untuk mendapatkan hasil yang optimal parameter-parameter tersebut perlu diatur nilainya. Untuk mengatur nilai parameter tersebut, pada penelitian ini akan menggunakan variasi *trial and error*. Untuk memperoleh hasil peramalan yang baik, akan dikombinasikan beberapa nilai *range* parameter, sehingga didapatkan nilai RMSE dan SMAPE yang kecil. Untuk mempermudah dalam pemilihan parameter maka terlebih dahulu akan dicari *range* parameter ε yang optimal, dimana untuk parameter C dan γ ditetapkan *range* nya. Langkah awal akan ditetapkan nilai *range* untuk parameter C adalah 10^{-2} -10 dan untuk parameter γ adalah 0,5-1,5. Sebagai percobaan akan digunakan model *input* $f(Y_{t-36})$ untuk mendapatkan parameter yang optimal. Berikut merupakan hasil percobaan nilai *range* parameter ε .

Tabel 4. 14 Hasil Percobaan *Range* Parameter ε Pos Cawak

Percobaan ke-	<i>Range</i> Parameter ε	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	0,0001-0,001	48,6128	66,7203	65,3783
2	0,001-0,01	48,5967	66,6940	65,3220
3	0,01-0,1	48,2775	66,2893	64,2731
4	0,1-0,5	47,9006	65,5899	62,1201
5	0,5-1	46,7701	61,1693	52,0880

Tabel 4.14 menunjukkan hasil bahwa *range* parameter ε yang optimal ketika nilai C dan γ tetap terdapat pada *range* 0,5-1, hal ini ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil dibandingkan pada *range* yang lain. Selanjutnya dengan cara yang sama akan dilakukan percobaan untuk mendapatkan *range* para-

meter γ yang optimal, dimana *range* parameter C adalah 10^{-2} - 10^4 dan untuk parameter ε adalah 0,5-1. Hasil percobaan beberapa *range* parameter γ terdapat pada Tabel 4.15.

Tabel 4. 15 Hasil Percobaan *Range* Parameter γ Pos Cawak

Percobaan ke-	<i>Range</i> Parameter γ	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	0,05-0,1	47,0327	62,3055	53,6263
2	0,1-0,5	46,9734	62,2725	53,5899
3	0,5-1	46,7701	61,1693	52,0880
4	0,6-1,5	46,7143	61,0097	51,8756
5	1-1,5	46,5133	60,5728	51,0380

Berdasarkan Tabel 4.15 didapatkan hasil bahwa *range* parameter γ yang optimal ketika nilai C dan ε tetap terdapat pada *range* 1-1,5 yang ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil. Langkah selanjutnya akan dilakukan percobaan untuk mendapatkan *range* parameter C yang optimal, dimana *range* untuk parameter ε dan γ didapatkan dari percobaan sebelumnya. Hasil beberapa percobaan *range* parameter C terdapat pada Tabel 4.16.

Tabel 4. 16 Hasil Percobaan *Range* Parameter C Pos Cawak

Percobaan ke-	<i>Range</i> Parameter C	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	0,01-10	46,5133	60,5728	51,0380
2	1-10	46,5109	60,5713	51,0417
3	5-50	46,3991	60,8767	51,0321
4	10-100	46,3549	60,8950	50,5822
5	100-150	46,3606	61,0868	50,8998

Dari Tabel 4.16 didapatkan informasi *range* parameter C yang optimal berada pada *range* 10-100. Dari beberapa percobaan yang telah dilakukan didapatkan hasil bahwa pada data curah hujan dasarian Pos Cawak, parameter C yang optimal pada *range* 10-100, parameter ε pada *range* 0,5-1 dan parameter γ pada *range* 1-1,5. Selanjutnya setelah didapatkan *range* parameter yang optimal, maka *range* tersebut akan digunakan untuk meramalkan curah hujan

dengan metode SVR yang selanjutnya akan dihitung kriteria kebaikan tiap model yang didapat seperti pada Tabel 4.17.

Tabel 4. 17 Kriteria Kebaikan Model SVR Pos Cawak

No	Model Input	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	$f(Y_{t-1}, Y_{t-2}, Y_{t-3})$	39,1760	69,8134	63,3755
2	$f(Y_{t-8}, Y_{t-10})$	48,1280	61,3600	62,6840
3	$f(Y_{t-10}, Y_{t-12}, Y_{t-35})$	41,5651	45,5568	53,0156
4	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-35})$	31,6725	66,1055	62,1517
5	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$	22,6186	53,4822	45,3539
6	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-39}, Y_{t-54}, Y_{t-55}, Y_{t-56}, Y_{t-57})$	22,9249	57,2007	47,7465
7	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-15}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-33}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-39}, Y_{t-51}, Y_{t-54}, Y_{t-55}, Y_{t-56}, Y_{t-57}, Y_{t-69})$	23,0598	58,5462	49,3260
8	$f(Y_{t-1}, Y_{t-36}, Y_{t-37})$	36,7956	61,3371	56,0077
9	$f(Y_{t-36})$	46,3806	60,8340	50,7196
10	$f(Y_{t-36})$	46,3806	60,8340	50,7196
11	$f(Y_{t-1}, Y_{t-36}, Y_{t-37}, Y_{t-72}, Y_{t-73})$	23,6894	65,7050	59,0880

Berdasarkan perhitungan kriteria *in sample* dan kriteria *out sample* RMSE dan SMAPE pada Tabel 4.17, didapatkan hasil yaitu model input $f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$ adalah model terbaik yang didapatkan dari metode SVR pada Pos Cawak yang ditunjukkan oleh nilai RMSE

dan SMAPE yang lebih kecil dari model SVR yang lain. Model peramalan SVR terbaik berdasarkan data *out sample* adalah

$$f(x_i) = (\mathbf{w} \Phi(x_i) + b) = \sum_{i,j=1}^n \alpha_i K(x_i, x_j) + b$$

$$f(x_i) = (\mathbf{w} \Phi(x_i) + b) = \sum_{i,j=1}^{537} \alpha_i \exp(-\gamma \|x_i - x_j\|^2) + b$$

dengan $x_j = [x_{j1}, x_{j2}, \dots, x_{j11}]$ dan $x_i = [x_{i1}, x_{i2}, \dots, x_{i11}]$, sehingga model terbaik SVR yang dihasilkan berdasarkan Tabel 4.17 sebagai berikut.

$$f(x_i) = \sum_{i,j=1}^{537} \alpha_i \exp(-1 \| [x_{i1}, x_{i2}, \dots, x_{i11}] - [x_{j1}, x_{j2}, \dots, x_{j11}] \|^2) + b$$

Model SVR tersebut memiliki 537 data *out sample*, dimana parameter $C=10$, $\gamma=1$ dan $\varepsilon=0,5$. *Support Vector* yang dihasilkan sebanyak 371, artinya terdapat 371 data yang taksiran parameter α_i tidak bernilai nol.

4.2.3 Peramalan Curah Hujan Menggunakan GA-SVR

Setelah dilakukan peramalan SVR, selanjutnya adalah mengoptimalkan parameter SVR menggunakan *Genetic Algorithm* (GA). Penggunaan *Genetic Algorithm* dimaksudkan untuk mengoptimalkan parameter SVR sehingga memiliki RMSE dan SMAPE yang lebih kecil. Pada *Genetic Algorithm* akan menggunakan 5 model SVR terbaik yang didapatkan pada analisis sebelumnya. Langkah awal yang dilakukan adalah menginisialisasi kromosom sebanyak 100. Nilai inisial kromosom akan menggunakan nilai parameter dari model SVR. Kromosom yang dibangkitkan memiliki tiga gen yang menunjukkan parameter dari model SVR dengan menggunakan fungsi kernel RBF yaitu *Cost* (C), γ dan ε . Sebagai ilustrasi akan digunakan model SVR dengan *input* $f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$, dimana nilai nilai parameter yang didapatkan dari SVR adalah $C=10$, $\gamma=1$ dan $\varepsilon=0,5$. Ilustrasi kromosom dengan tiga gen untuk Pos Cawak terdapat pada Gambar 4.13

Parameter	$Cost$	γ	ε
Kromosom	10	1	0,5

Gambar 4. 13 Ilustrasi Kromosom Parameter SVR dengan Tiga Gen Untuk Pos Cawak

Gambar 4.13 menunjukkan ilustrasi kromosom dengan tiga gen yang diperoleh dari model SVR. Nilai parameter yang sebelumnya didapatkan dari SVR akan digunakan untuk nilai inisial pada GA-SVR. Kromosom yang terbentuk akan menjalani proses seleksi *roulette wheel*, *crossover*, *mutation* dan *elitism* hingga diperoleh hasil yang konvergen dengan nilai RMSE dan SMAPE yang lebih kecil.

Sebelum melewati tahapan *Genetic Algorithm* terlebih dahulu ditentukan nilai *fitness*. Nilai *fitness* merupakan acuan dalam tahapan *Genetic Algorithm* seperti melakukan proses seleksi, pin-dah silang (*crossover*), mutasi dan *elitism* sehingga didapatkan nilai *fitness* terbaik. Nilai *fitness* yang digunakan adalah nilai kesalahan dalam peramalan yaitu RMSE. Fungsi objektif yang ingin dicapai adalah untuk meminimumkan nilai RMSE pada model SVR. Sebelum melewati tahapan dalam *Genetic Algorithm* maka akan dihitung terlebih dahulu nilai *fitness* berdasarkan kromosom-kromosom yang terbentuk. Pada ilustrasi akan dicontohkan model optimasi parameter SVR pada model *input* $f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$. Ilustrasi nilai *fitness* masing-masing kromosom dapat dilihat pada Tabel 4.18.

Tabel 4. 18 Ilustrasi Nilai *Fitness* Setiap Kromosom Pos Cawak

Kromosom ke-	Kromosom			Nilai <i>Fitness</i>
	C	γ	ε	
1	9,9987	1,0012	0,4998	22,6109
2	10,0076	1,0152	0,4876	22,1035
:	:	:	:	:
100	9,5634	0.9985	0,4765	21,6262

Berdasarkan Tabel 4.18 diperoleh nilai *fitness* yang memiliki nilai berbeda-beda pada tiap kromosom. Nilai *fitness* tersebut akan menjadi acuan dalam melakukan proses *Genetic Algorithm* selanjutnya hingga diperoleh nilai *fitness* yang paling minimum. Tahap selanjutnya adalah seleksi, dimana terdapat beberapa metode seleksi yang dapat digunakan. Penelitian ini akan menggunakan metode seleksi *roulette wheel*. Seleksi *roulette wheel* merupakan salah satu cara dalam penentuan orang tua kromosom yang akan melakukan *crossover* atau pindah silang. Orang tua yang terpilih dalam proses *roulette wheel* melibatkan nilai *fitness* relatif dan *fitness* kumulatif. Ilustrasi nilai *fitness* relatif dan *fitness* kumulatif terdapat pada Tabel 4.19.

Tabel 4. 19 Ilustrasi Nilai *Fitness* Relatif dan *Fitness* Kumulatif Pos Cawak

Kromosom ke-	<i>Fitness</i>	<i>Fitness Relatif</i>	<i>Fitness Kumulatif</i>
1	22,6109	0,0097	0,0097
2	22,1035	0,0096	0,0195
:	:	:	:
100	21,6262	0,0094	1

Setelah diperoleh nilai *fitness* relatif dan *fitness* kumulatif, selanjutnya adalah membangkitkan bilangan *random* antara 0 dan 1 sebanyak jumlah kromosom yaitu 100. Bilangan *random* ini digunakan sebagai pembanding untuk setiap individu apakah individu tersebut layak menjadi individu baru atau tidak. Tabel 4.20 menunjukkan ilustrasi perbandingan bilangan *random* dengan nilai *fitness* kumulatif yang akan digunakan pada tahap seleksi *roulette wheel*.

Tabel 4.20 menunjukkan bahwa nilai bilangan random kromosom pertama adalah 0,0160. Nilai ini lebih kecil jika dibandingkan dengan nilai *fitness* kumulatif kromosom kedua, sehingga kromosom ke-2 akan menjadi calon orang tua untuk generasi selanjutnya. Pada kromosom kedua memiliki bilangan random kurang dari nilai *fitness* kumulatif kromosom ke-1 sehingga

kromosom ke-1 terpilih untuk calon orang tua untuk generasi selanjutnya.

Tabel 4. 20 Ilustrasi Nilai Perbandingan *Fitness* Kumulatif dan Bilangan Random Pos Cawak

Kromosom ke-	<i>Fitness</i> Kumulatif	Bilangan Random
1	0,0097	0,0160
2	0,0195	0,0023
:	:	:
100	1	0,9765

Proses selanjutnya adalah *crossover* atau pindah silang. *Crossover* melibatkan dua kromosom orang tua yang akan membentuk dua kromosom anak. Jika bilangan *random* lebih kecil dari probabilitas pindah silang ($P_c = 0,8$), maka orang tua yang terpilih dari individu satu dan individu dua akan dikawinkan dan keturunannya akan menjadi individu baru. Tipe *crossover* yang digunakan dalam penelitian ini adalah *local arithmetic crossover*. Tipe *crossover* ini banyak digunakan untuk kasus yang menggunakan nilai bilangan *real* (*real values encoding*) (Eiben, 2003). Berikut algoritma perhitungan *local arithmetic crossover*.

$$C_i^{gen+1} = \alpha C_i^{gen} + (1 - \alpha) C_j^{gen}$$

$$C_j^{gen+1} = \alpha C_j^{gen} + (1 - \alpha) C_i^{gen}$$

dimana C^{gen} adalah kromosom orang tua, C^{gen+1} adalah kromosom anak hasil *crossover* dan α adalah bobot, dimana nilai α diantara *range* 0 dan 1. Gambar 4.14 menunjukkan ilustrasi proses *crossover* kromosom orang tua satu dan kromosom orangtua dua yang akan menghasilkan anak 1 dan anak 2.

Gambar 4.14 menunjukkan ilustrasi kromosom orangtua dan kromosom hasil *crossover*. Sebagai ilustrasi misalkan diperoleh nilai $\alpha = 0,6078$, kemudian akan dihitung berdasarkan rumus sebelumnya, sehingga dari *crossover* antara dua orang tua akan menghasilkan anak 1 dan anak 2. Proses selanjutnya adalah proses

mutasi. Proses memilih gen yang akan mengalami mutasi dilakukan dengan membangkitkan bilangan *random* antara 0 sampai 1 sebanyak gen dalam kromosom. Apabila bilangan *random* lebih kecil dari probabilitas mutasi, $P_m = 0,01$ maka gen tersebut mengalami mutasi dengan cara mengganti gen tersebut dengan bilangan acak. Ilustrasi proses mutasi pada Gambar 4.15.

Orang tua 1	10,0087	1,0123	0,4998
Orang tua 2	9,9988	1,0012	0,4765
Anak 1	10,0048	1,0079	0,4907
Anak 2	9,9988	1,0012	0,4765

Gambar 4. 14 Ilustrasi Kromosom yang Melakukan *Crossover* Pos Cawak

Sebelum dimutasi	Bilangan Random	0,5361	0,0032	0,2342
	Kromosom	9,9988	1,0123	0,4998
Setelah dimutasi	Kromosom	9,9988	1,0128	0,4998

Gambar 4. 15 Ilustrasi Proses Mutasi Pos Cawak

Proses mutasi bertujuan untuk mencegah terjadinya *convergence premature*. Gambar 4.15 menunjukkan bahwa gen ke-2 adalah gen yang dilakukan mutasi, karena pada gen tersebut memiliki bilangan random yang lebih kecil dari P_m , sehingga gen dalam kromosom tersebut akan diganti dengan bilangan acak.

Proses terakhir pada GA yang digunakan untuk memperoleh nilai *fitness* terbaik adalah *elitism*. Proses *elitism* bertujuan untuk mempertahankan estimasi parameter yang menghasilkan nilai *fitness* terkecil untuk generasi selanjutnya. Jumlah kromosom terbaik

yang dipertahankan untuk generasi selanjutnya adalah adalah 5 kromosom. Berikut disajikan ilustrasi proses *elitism*.

Tabel 4. 21 Ilustrasi Proses *Elitism* Generasi Pertama Pos Cawak

Kromosom ke-	Kromosom			Nilai <i>Fitness</i>
	C	γ	ϵ	
1	10,0001	1,0128	0,4998	21,4567
2	9,99982	1,0098	0,4768	21,5634
:	:	:	:	:
100	9,7656	0,9965	0,4987	22,0098

Kromosom yang bertahan ke generasi selanjutnya adalah kromosom yang menghasilkan *fitness* terkecil. Lima kromosom dengan nilai *fitness* terbaik akan bertahan pada generasi kedua. Proses *Elitism* generasi kedua ditunjukkan pada Tabel 4.22.

Tabel 4. 22 Ilustrasi Proses *Elitism* Generasi Kedua Pos Cawak

Kromosom ke-	Kromosom			Nilai <i>Fitness</i>
	C	γ	ϵ	
1	10,0001	1,0128	0,4998	21,4567
2	9,99982	1,0098	0,4768	21,5634
:	:	:	:	:
100	9,2345	0,9987	0,4876	22,0009

Proses selanjutnya generasi kedua juga akan melakukan proses *elitism* hingga generasi-generasi selanjutnya sampai kondisi *fitness* telah konvergen. Hasil optimasi 5 model SVR terbaik yang telah didapatkan pada analisis sebelumnya ditunjukkan pada Tabel 4.23.

Berdasarkan Tabel 4.23 didapatkan informasi bahwa dari 5 model SVR terbaik, maka model *input* $f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-2}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$ adalah model terbaik yang didapatkan dari metode GA-SVR. Model *input* ini memiliki nilai RMSE dan SMAPE yang lebih kecil dibandingkan dengan model GA-SVR yang lain.

Tabel 4. 23 Hasil Optimasi Parameter SVR Pos Cawak Menggunakan GA

No	Model Input	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	$f(Y_{t-10}, Y_{t-12}, Y_{t-35})$	41,4649	45,3635	53,2349
2	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-2}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$	18,3483	53,4608	45,0682
3	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-39}, Y_{t-54}, Y_{t-55}, Y_{t-56}, Y_{t-57})$	21,6371	56,9737	47,3214
4	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-15}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-33}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-39}, Y_{t-51}, Y_{t-54}, Y_{t-55}, Y_{t-56}, Y_{t-57}, Y_{t-69})$	23,0516	58,5363	49,3086
5	$fY_{t-36})$	46,3432	60,4113	50,1810

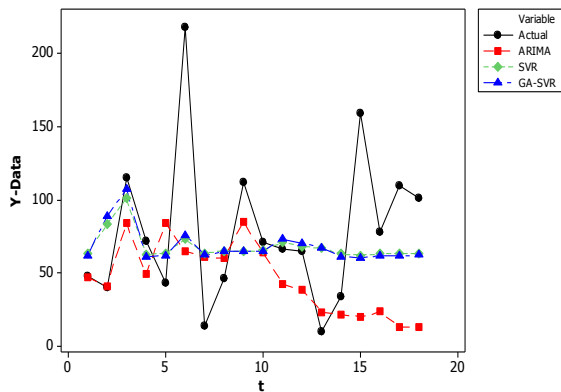
4.2.4 Perbandingan Model Terbaik Pos Cawak

Langkah selanjutnya adalah melakukan perbandingan performansi model terbaik berdasarkan nilai RMSE *in sample*, RMSE *out sample* dan SMAPE *out sample*. Hasil perbandingan model peramalan terbaik ditunjukkan pada Tabel 4.24.

Tabel 4. 24 Hasil Perbandingan Metode Peramalan Pos Cawak

Metode	Model	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
ARIMA	ARIMA (0,0,1)(0,1,1) ³⁶	49,9171	62,7889	68,3422
SVR	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$	22,6186	53,4822	45,3539
GA-SVR	$f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$	18,3483	53,4608	45,0682

Berdasarkan Tabel 4.24 diketahui bahwa model peramalan terbaik untuk meramalkan curah hujan Pos Cawak adalah model $input f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$ yang didapatkan dari metode GA-SVR. Model ini memiliki ketepatan akurasi yang paling tinggi dibandingkan metode lainnya. Gambar 4.16 menunjukkan perbandingan secara visual pada data *out sample* dari ketiga metode peramalan. Setelah diperoleh model terbaik, maka di-lakukan *forecast* curah hujan 18 periode ke depan yakni pada bulan Juli 2016 sampai dengan Desember 2016 menggunakan metode GA-SVR yang terdapat pada Tabel 4.25.



Gambar 4. 16 Perbandingan Visual pada Data *Out Sample* Pos Cawak

Tabel 4. 25 Hasil *Forecast* Curah Hujan 6 Periode Selanjutnya Pos Cawak

Dasarian	Bulan	Forecast
1	Juli	52,1183
2	Juli	50,3830
3	Juli	46,2815
1	Agustus	39,4695
2	Agustus	41,1700
3	Agustus	48,0710

Tabel 4. 25 Hasil *Forecast* Curah Hujan 6 Periode Selanjutnya Untuk Pos Cawak (Lanjutan)

Dasarian	Bulan	Forecast
1	September	48,7525
2	September	47,8739
3	September	49,7663
1	Oktober	40,4566
2	Oktober	40,3609
3	Oktober	30,8328
1	November	50,8710
2	November	55,0822
3	November	51,2971
1	Desember	51,7623
2	Desember	51,6139
3	Desember	51,2309

4.2.5 Kalender Tanam Padi Pos Cawak Periode Juli - Desember Tahun 2016

Berdasarkan hasil *forecast* curah hujan dasarian pada Pos Cawak pada bulan Juli hingga bulan Desember 2016, diketahui bahwa curah hujan selama 6 bulan tersebut antara 50-100 mm dan kurang dari 50 mm. Curah hujan ini termasuk pada kategori hujan sedang dan lebat (Suriadikusumah, 2007). Berikut adalah susunan kalender tanam padi untuk wilayah Cawak Kabupaten Bojonegoro bulan Juli 2016- Desember 2016.

Juli			Agustus			September			Oktober			November			Desember		
1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3



Musim tanam palawija

Musim tanam padi dengan bantuan pengairan

Gambar 4. 17 Kalender Tanam Padi Pos Cawak Periode Juli 2016 hingga Desember 2016

Berdasarkan Gambar 4.17, kebutuhan air untuk penanaman padi sepanjang bulan Juli 2016- Desember 2016 tidak terpenuhi, karena penanaman padi sawah membutuhkan air rata-rata 145 mm pada musim hujan (Oldeman, 1980), sehingga penanaman padi harus dibantu dengan pengairan dan tidak mengandalkan turunnya hujan. Pada bulan Agustus hingga Oktober petani dapat mengganti padi dengan palawija, dikarenakan palawija membutuhkan air yang lebih sedikit sekitar 50 mm pada musim kemarau (Oldeman, 1980). Jika persediaan cadangan air terbatas, maka petani dapat melakukan tanam palawija untuk mengantisipasi kekeringan atau kekurangan air. Namun, jika dipaksakan menanam padi, maka petani dan pemerintah harus memastikan tersedianya cadangan air dari embung atau waduk didaerah sekitar Pos Cawak.

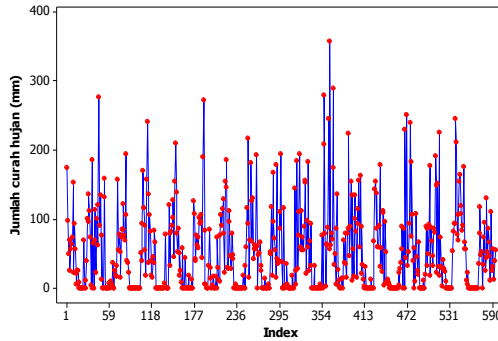
4.3 Peramalan Curah Hujan Pos Kedungadem

Peramalan curah hujan di Pos Kedungadem dilakukan menggunakan tiga metode yaitu ARIMA, SVR dan GA-SVR. Peramalan dengan metode SVR dan GA-SVR akan didasarkan pada *input* dari model ARIMA yang telah memiliki parameter yang signifikan. Setelah didapatkan model terbaik akan dilakukan pemilihan metode terbaik berdasarkan dengan membandingkan nilai RMSE dan SMAPE.

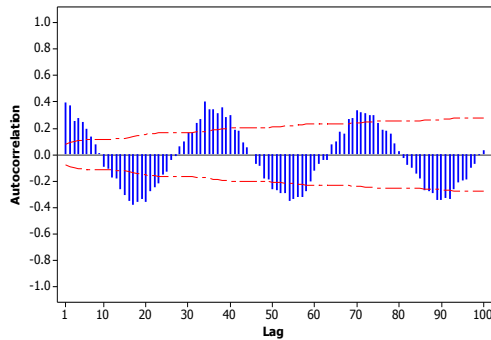
4.3.1 Peramalan Curah Hujan Menggunakan ARIMA

Langkah pertama yang harus dilakukan pada peramalan dengan ARIMA adalah identifikasi terhadap pola data untuk mengetahui apakah data curah hujan dasarian Pos Kedungadem sudah memenuhi asumsi stasioner terhadap *mean* dan varians. Dalam menentukan model peramalan curah hujan di Pos Kedungadem menggunakan metode ARIMA, perlu melihat *time series plot* dari data untuk mengetahui pola dan karakteristik dari data.

Berdasarkan Gambar 4.18 menunjukkan bahwa curah hujan dasarian di Pos Kedungadem belum stasioner terhadap *mean* dan varians dikarenakan fluktuasi antara titik satu dengan titik lainnya sangat bervariasi. Plot ACF pada Gambar 4.19 menunjukkan data memiliki pola musiman (*seasonal*).

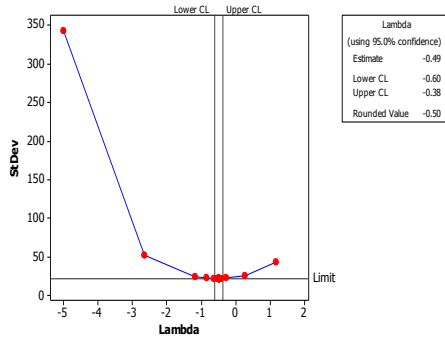


Gambar 4. 18 Time Series Plot Curah Hujan Dasarian Pos Kedungadem



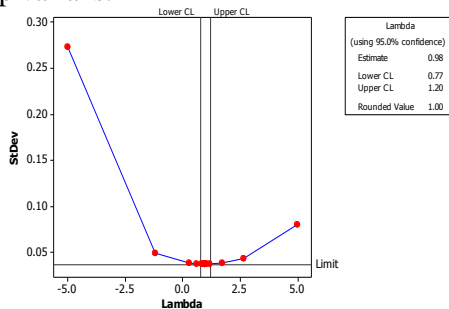
Gambar 4. 19 Plot ACF Curah Hujan Dasarian Pos Kedungadem

Langkah selanjutnya adalah dilakukan analisis dengan melihat plot *Box-Cox* untuk mengetahui apakah data sudah stasioner terhadap varians seperti pada Gambar 4.20. Gambar 4.20 menunjukkan bahwa data curah hujan dasarian di Pos Kedungadem belum stasioner terhadap varians. Hal ini ditunjukkan dengan nilai *Lower CL* dan *Upper CL* yang belum melewati angka 1 atau bisa juga dilihat pada nilai *rounded value* (λ) yang belum bernilai satu. Karena nilai λ sebesar -0,5 maka data akan ditransformasikan $1/\sqrt{Y_t}$ untuk menstabilkan varians.



Gambar 4. 20 *Box-Cox* Data Curah Hujan Dasarian Pos Kedungadem

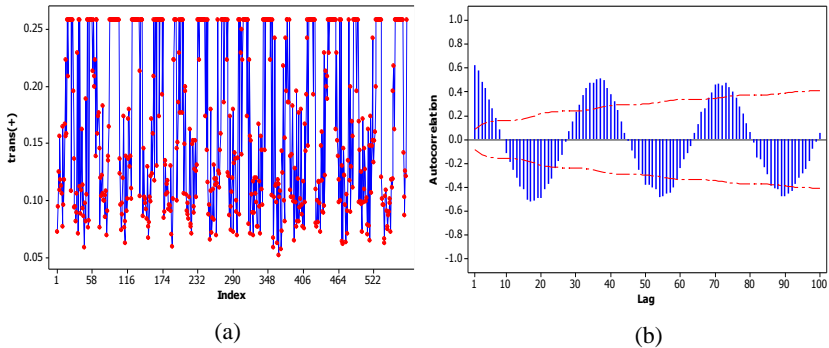
Setelah dilakukan transformasi $1/\sqrt{Y_t}$ maka dilihat kembali pada plot *Box-Cox* untuk mengetahui apakah data sudah stasioner terhadap varians atau belum. Gambar 4.21 merupakan plot *Box-Cox* curah hujan dasarian di Pos Kedungadem yang telah dilakukan transformasi. Dapat dilihat bahwa nilai *Lower CL* dan *Upper CL* yang sudah melewati angka 1 dan nilai *rounded value* (λ) yang bernilai satu, sehingga data curah hujan hasil transformasi sudah stasioner terhadap varians.



Gambar 4. 21 *Box-Cox* Data Curah Hujan Dasarian Pos Kedungadem Hasil Transformasi

Setelah stasioneritas data terhadap varians terpenuhi, maka dilanjutkan dengan pengujian stasioneritas data terhadap *mean* menggunakan Uji *Augmented Dickey Fuller* (ADF). Uji ADF dilakukan dalam beberapa *lag* yakni *lag* 1, *lag* 2, *lag* 5, *lag* 8, *lag* 15

dan *lag* 30, dimana dari beberapa *lag* tersebut didapatkan hasil *p-value* yang sama yakni sebesar 0,01. Nilai ini menunjukkan *P-value* $< \alpha$ ($\alpha=0,05$), sehingga tolak H_0 maka disimpulkan bahwa data transformasi sudah stasioner terhadap *mean*. *Time Series Plot* dan plot ACF data Transformasi terdapat pada Gambar 4.22.



Gambar 4. 22 *Time Series Plot* (a) dan plot ACf (b) Data Hasil Transformasi Pos Kedungadem

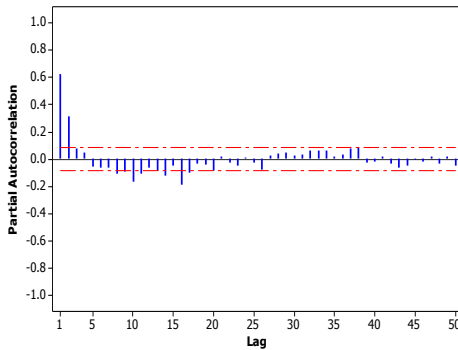
Gambar 4.22 (a) menunjukkan data curah hujan hasil transformasi secara visual belum stasioner terhadap *mean* yang ditunjukkan plot yang tidak berfluktuasi disekitar garis *mean* dan plot ACF menunjukkan data memiliki pola musiman (*seasonal*). Jika berdasarkan pada uji ADF maka data transformasi telah stasioner terhadap *mean*, namun hasil ini berbeda dengan *time series plot*. Oleh karena itu, untuk analisis ARIMA selanjutnya pada data hasil transformasi akan dibagi menjadi tanpa *differencing*, *differencing* musiman 18 dan *differencing* musiman 36.

A. ARIMA Pos Kedungadem Tanpa Differencing

Setelah asumsi stasioneritas data terhadap *mean* dan varians terpenuhi, maka selanjutnya dilakukan identifikasi model ARIMA sementara berdasarkan pada plot ACF dan PACF data hasil transformasi sebagai berikut.

Gambar 4.22 (b) menunjukkan bahwa plot ACF memiliki pola *dies down*. Pada plot PACF pada Gambar 4.23 diketahui ter-

dapat beberapa *lag* yang keluar dari batas *significance limit* yang menunjukkan *lag* tersebut signifikan. *Lag* tersebut adalah *lag* 1, *lag* 2, *lag* 8, *lag* 9, *lag* 10, *lag* 11, *lag* 14, *lag* 16, *lag* 17 *lag* 20, *lag* 37 dan *lag* 38. Berdasarkan plot ACF dan plot PACF diperoleh beberapa model ARIMA sementara yaitu ARIMA (2,0,0), ARIMA ([1,2,8], 0,0), ARIMA([1,8],0,0), ARIMA ([1,2,17],0,0), ARIMA ([1,2,10,16],0,0) dan ARIMA ([8,10],0,0).



Gambar 4. 23 Plot PACF Data Transformasi Curah Hujan Pos Kedungadem

Setelah menetapkan beberapa model dugaan ARIMA, maka langkah selanjutnya adalah mengestimasi parameter-parameter model dengan metode *Conditional Least Square* (CLS). Parameter-parameter tersebut akan di uji signifikansinya menggunakan uji *t*. Berikut adalah estimasi parameter dan pengujian signifikansi parameter beberapa model dugaan ARIMA di Pos Kedungadem tanpa *differencing*.

Tabel 4. 26 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem Tanpa *Differencing*

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (2,0,0)	$\hat{\phi}_1 = 0,5411$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,4264$	<0,0001	Signifikan
ARIMA ([1,2,8],0,0)	$\hat{\phi}_1 = 0,5261$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,4015$	<0,0001	Signifikan
	$\hat{\phi}_8 = 0,0445$	0,1152	Tidak Signifikan

Tabel 4. 26 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem Tanpa *Differencing* (Lanjutan)

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA ([1,2,17],0,0)	$\hat{\phi}_1 = 0,5307$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,4049$	<0,0001	Signifikan
	$\hat{\phi}_{17} = 0,0417$	0,0437	Signifikan
ARIMA ([1,2,10,16],0,0)	$\hat{\phi}_1 = 0,5337$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,4191$	<0,0001	Signifikan
	$\hat{\phi}_{10} = -0,0106$	0,7544	Tidak Signifikan
	$\hat{\phi}_{16} = -0,0307$	0,2725	Tidak Signifikan
ARIMA ([8,10],0,0)	$\hat{\phi}_8 = 0,6902$	<0,0001	Signifikan
	$\hat{\phi}_{10} = 0,1819$	0,0031	Signifikan
ARIMA ([1,8],0,0)	$\hat{\phi}_1 = 0,8123$	<0,0001	Signifikan
	$\hat{\phi}_8 = 0,1503$	<0,0001	Signifikan

Berdasarkan Tabel 4.26 didapatkan informasi mengenai hasil estimasi dan pengujian signifikansi parameter model dugaan ARIMA di Pos Kedungadem tanpa *differencing*. Tabel 4.26 menunjukkan bahwa terdapat 2 model yang memiliki parameter yang tidak signifikan yang ditunjukkan oleh nilai *P-value* yang lebih dari α ($\alpha=0,05$), sehingga gagal tolak H_0 . Dapat disimpulkan dari 6 model dugaan ARIMA terdapat 4 model yang memiliki parameter yang sudah signifikan yaitu ARIMA (2,0,0), ARIMA ([1,8],0,0), ARIMA ([1,2,17],0,0) dan ARIMA ([8,10],0,0).

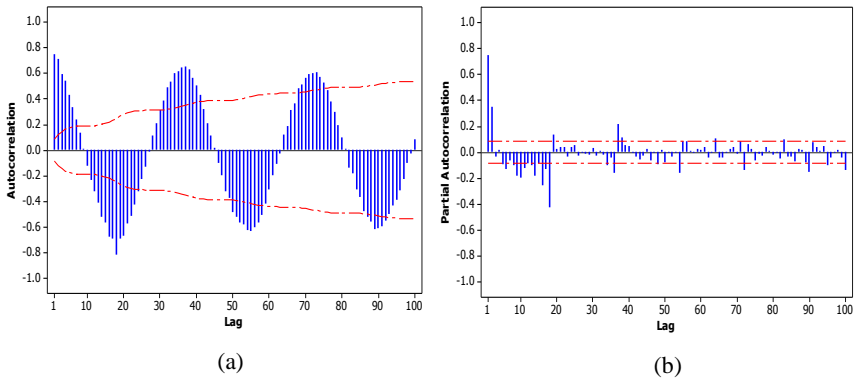
Langkah selanjutnya akan dilakukan *diagnostic checking* untuk mengetahui apakah residual sudah memenuhi asumsi *white noise* dan berdistribusi normal. Hasil *diagnostic checking* terdapat pada Tabel 4.27. Tabel 4.27 menunjukkan bahwa tidak terdapat model dugaan ARIMA yang dapat memenuhi asumsi *white noise*. Hal ini ditunjukkan oleh nilai *P-value* yang kurang dari α ($\alpha=0,05$), sehingga tolak H_0 . Dikarenakan tidak terdapat model ARIMA yang memenuhi asumsi, maka analisis selanjutnya akan dilakukan *differencing* musiman pada data hasil transformasi.

Tabel 4. 27 *Diagnostic Checking Model ARIMA Pos Kedungadem Tanpa Differencing*

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (2,0,0)	6	0,0020	<0,0100
	12	0,0103	
	18	0,0017	
	24	0,0014	
	30	0,0027	
	36	0,0010	
	42	<0,0001	
	48	0,0002	
ARIMA ([1,2,17],0,0)	6	0,0025	<0,0100
	12	0,0125	
	18	0,0002	
	24	<0,0001	
	30	0,0002	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	
ARIMA ([8,10],0,0)	6	<0,0001	<0,0100
	12	<0,0001	
	18	<0,0001	
	24	<0,0001	
	30	<0,0001	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	
ARIMA ([1,8],0,0)	6	<0,0001	<0,0100
	12	<0,0001	
	18	<0,0001	
	24	<0,0001	
	30	<0,0001	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	

B. ARIMA Pos Kedungadem dengan *Differencing* Musiman 18

Gambar 4.22 (a) menunjukkan bahwa *Time series plot* data transformasi belum stasioner terhadap *mean*, serta plot ACF data transformasi pada Gambar 4.22 (b) diketahui bahwa data memiliki pola musiman. Oleh karena itu akan dilakukan *differencing* musiman pada *lag* 18 terhadap data hasil transformasi. *Lag* tersebut dipilih dikarenakan diduga data curah hujan dasarian musiman selama 6 bulan. Plot ACF dan plot PACF data transformasi setelah dilakukan *differencing* musiman 18 terdapat pada Gambar 4.24.



Gambar 4. 24 Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah *Differencing* Musiman 18 Pos Kedungadem

Gambar 4.24 (a) dan Gambar 4.24 (b) menunjukkan plot ACF dan PACF data hasil *differencing* musiman 18 dari data transformasi. Berdasarkan plot PACF dan ACF dapat ditentukan orde dari model ARIMA sementara. Pada plot ACF menunjukkan pola *dies down*. Berdasarkan plot PACF didapatkan beberapa *lag* yang keluar dari batas *significance limit* yaitu pada *lag* 1, *lag* 2, *lag* 6, *lag* 9, *lag* 10, *lag* 14, *lag* 16, *lag* 18, *lag* 19, *lag* 36, *lag* 37, *lag* 54, *lag* 72, *lag* 90, dan *lag* 100. Berdasarkan plot ACF dan PACF didapatkan beberapa model ARIMA sementara hasil *differencing* musiman 18 yaitu ARIMA (2,0,0) (2,1,0)¹⁸, ARIMA ([1,2,16],0,0)

$(3,1,0)^{18}$, ARIMA $([1,16,37],0,0)$ $(3,1,0)^{18}$ dan ARIMA $([1,2,6],0,0)$ $(3,1,0)^{18}$.

Setelah menetapkan beberapa model dugaan ARIMA, maka selanjutnya dilakukan estimasi parameter model dengan metode *Conditional Least Square* (CLS). Kemudian parameter-parameter tersebut akan di uji signifikansinya menggunakan uji *t*. Berikut adalah estimasi dan pengujian signifikansi parameter beberapa model dugaan ARIMA di Pos Kedungadem.

Tabel 4. 28 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem dengan *Differencing* Musiman 18

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (2,0,0) (2,1,0) ¹⁸	$\hat{\phi}_1 = 0,3821$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,2547$	<0,0001	Signifikan
	$\hat{\Phi}_{18} = -0,7181$	<0,0001	Signifikan
	$\hat{\Phi}_{36} = -0,1652$	0,0001	Signifikan
ARIMA ([1,2,16],0,0) (3,1,0) ¹⁸	$\hat{\phi}_1 = 0,3301$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,2312$	<0,0001	Signifikan
	$\hat{\phi}_{16} = -0,2826$	<0,0001	Signifikan
	$\hat{\Phi}_{18} = -0,6863$	<0,0001	Signifikan
	$\hat{\Phi}_{36} = -0,4055$	<0,0001	Signifikan
ARIMA ([1,16,37],0,0) (3,1,0) ¹⁸	$\hat{\Phi}_{54} = -0,2460$	<0,0001	Signifikan
	$\hat{\phi}_1 = 0,3865$	<0,0001	Signifikan
	$\hat{\phi}_{16} = -0,2616$	<0,0001	Signifikan
	$\hat{\phi}_{37} = 0,2108$	<0,0001	Signifikan
	$\hat{\Phi}_{18} = -0,7179$	<0,0001	Signifikan
	$\hat{\Phi}_{36} = -0,4020$	<0,0001	Signifikan
ARIMA ([1,2,6],0,0)(3,1,0) ¹⁸	$\hat{\Phi}_{54} = -0,2191$	<0,0001	Signifikan
	$\hat{\phi}_1 = 0,2726$	<0,0001	Signifikan
	$\hat{\phi}_2 = 0,1729$	<0,0001	Signifikan
	$\hat{\phi}_6 = 0,0169$	0,6769	Tidak Signifikan
	$\hat{\Phi}_{18} = -0,8315$	<0,0001	Signifikan
	$\hat{\Phi}_{36} = -0,3904$	<0,0001	Signifikan
	$\hat{\Phi}_{54} = -0,3594$	<0,0001	Signifikan

Tabel 4.28 menunjukkan terdapat satu model yang memiliki parameter yang tidak signifikan yang ditunjukkan oleh *P-value* le-

bih dari α ($\alpha=0,05$). Dapat disimpulkan dari 4 model dugaan ARIMA terdapat 3 model yang memiliki parameter yang signifikan yaitu ARIMA (2,0,0) (2,1,0)¹⁸, ARIMA ([1,2,16],0,0)(3,1,0)¹⁸ dan ARIMA ([1,16,37],0,0) (3,1,0)¹⁸.

Langkah selanjutnya adalah *diagnostic checking* untuk mengetahui apakah residual sudah memenuhi asumsi *white noise* dan berdistribusi normal seperti pada Tabel 4.29, Tabel 4.29 menunjukkan tidak terdapat model ARIMA Pos Kedungadem dengan *differencing* musiman 18 yang dapat memenuhi asumsi residual *white noise*. Hal ini ditunjukkan dengan nilai *P-value* yang kurang dari α ($\alpha=0,05$), sehingga tolak H_0 . Untuk analisis selanjutnya akan digunakan *differencing* musiman 36 terhadap data transformasi.

Tabel 4. 29 *Diagnostic Checking Model ARIMA Pos Kedungadem dengan Differencing Musiman 18*

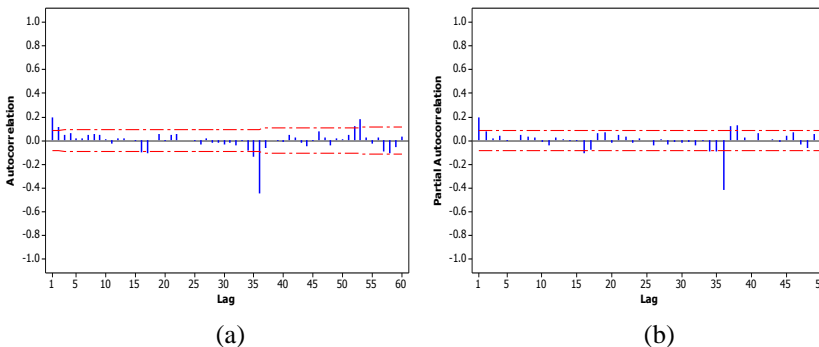
Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (2,0,0) (2,1,0) ¹⁸	6	0,0015	>0,1500
	12	0,0145	
	18	<0,0001	
	24	<0,0001	
	30	<0,0001	
	36	<0,0001	
	42	<0,0001	
	48	<0,0001	
ARIMA ([1,2,16],0,0) (3,1,0) ¹⁸	12	0,0233	>0,1500
	18	0,0518	
	24	0,0279	
	30	0,0394	
	36	0,0712	
	42	0,0223	
	48	0,0398	

Tabel 4. 29 *Diagnostic Checking Model ARIMA Pos Kedungadem dengan Differencing Musiman 18 (Lanjutan)*

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA ([1,16,37],0,0) (3,1,0) ¹⁸	12	0,0001	>0,1500
	18	0,0027	
	24	0,0028	
	30	0,0142	
	36	0,0245	
	42	0,0312	
	48	0,0454	

C. ARIMA Pos Kedungadem dengan *Differencing* Musiman 36

Time series plot pada Gambar 4.22 (a) menunjukkan bahwa data hasil transformasi belum stasioner terhadap *mean*, serta plot ACF data tranformasi pada Gambar 4.22 (b) diketahui bahwa data memiliki pola musiman (*seasonal*). Oleh karena itu akan dilakukan *differencing* musiman pada *lag* 36 terhadap data hasil transformasi. *Lag* tersebut dipilih dikarenakan diduga data curah hujan dasarian musiman selama 12 bulan. Gambar 4.25 menunjukkan plot ACF dan PACF data transformasi setelah dilakukan *differencing* musiman 36.



Gambar 4. 25 Plot ACF (a) dan Plot PACF (b) Data Transformasi Curah Hujan Setelah *Differencing* Musiman 36 Pos Kedungadem

Berdasarkan plot ACF pada Gambar 4.25 (a) terlihat bahwa *lag 1, lag 2, lag 17, lag 35, lag 36, lag 52 dan lag 53* adalah *lag* yang keluar dari batas *significance limit*. Plot PACF pada Gambar 4.25 (b) juga menunjukkan terdapat beberapa *lag* yang keluar dari batas *significance limit* yaitu *lag 1, lag 16, lag 36, lag 37 dan lag 38*. Beberapa model ARIMA sementara berdasarkan plot ACF dan PACF hasil *differencing* musiman 36 yaitu ARIMA (0,0,1) (0,1,1)³⁶, ARIMA (0,0,2) (0,1,1)³⁶, ARIMA ([1,16,37],0,0) (1,1,0)³⁶, ARIMA (1,0,1)(0,1,1)³⁶ dan ARIMA(1,0,[1,52]) (0,1,1)³⁶.

Setelah menetapkan beberapa model dugaan ARIMA, maka langkah selanjutnya adalah mengestimasi parameter model dengan metode *Conditional Least Square* (CLS). Kemudian parameter tersebut akan di uji signifikansinya dengan uji *t*. Berikut adalah estimasi dan pengujian signifikansi parameter beberapa model dugaan ARIMA di Pos Kedungadem.

Tabel 4. 30 Estimasi dan Pengujian Signifikansi Parameter Model Dugaan ARIMA Pos Kedungadem dengan *Differencing* Musiman 36

Model ARIMA	Parameter	P-Value	Keputusan
ARIMA (0,0,1) (0,1,1) ³⁶	$\hat{\theta}_1 = -0,1385$	0,0014	Signifikan
	$\hat{\theta}_{36} = 0,7274$	<0,0001	Signifikan
ARIMA (0,0,2) (0,1,1) ³⁶	$\hat{\theta}_1 = -0,1441$	<0,0001	Signifikan
	$\hat{\theta}_2 = -0,1210$	0,0051	Signifikan
	$\hat{\theta}_{36} = 0,7317$	<0,0001	Signifikan
ARIMA (1,0,1)(0,1,1) ³⁶	$\hat{\theta}_1 = 0,5319$	0,0005	Signifikan
	$\hat{\theta}_{36} = 0,7296$	<0,0001	Signifikan
	$\hat{\phi}_1 = 0,6870$	<0,0001	Signifikan
ARIMA ([1,16,37],0,0) (1,1,0) ³⁶	$\hat{\phi}_1 = 0,1923$	<0,0001	Signifikan
	$\hat{\phi}_{16} = -0,0931$	0,0300	Signifikan
	$\hat{\phi}_{37} = 0,0243$	0,5846	Tidak Signifikan
	$\hat{\Phi}_{36} = -0,4630$	<0,0001	Signifikan
ARIMA (1,0,[1,52]) (0,1,1) ³⁶	$\hat{\theta}_1 = 0,5322$	<0,0001	Signifikan
	$\hat{\theta}_{52} = -0,0903$	0,0194	Signifikan
	$\hat{\theta}_{36} = 0,7259$	<0,0001	Signifikan
	$\hat{\phi}_1 = 0,6886$	<0,0001	Signifikan

Berdasarkan Tabel 4.30 diketahui terdapat satu model yang memiliki parameter yang tidak signifikan yang ditunjukkan oleh nilai P -value yang lebih dari α ($\alpha=0,05$) sehingga gagal tolak H_0 . Dapat disimpulkan dari 5 model dugaan ARIMA terdapat 4 model yang memiliki parameter yang signifikan yaitu ARIMA (0,0,1) (0,1,1)³⁶, ARIMA (0,0,2)(0,1,1)³⁶, ARIMA(1,0,1)(0,1,1)³⁶ dan ARIMA (1,0,[1,52]) (0,1,1)³⁶.

Langkah selanjutnya adalah *diagnostic checking* untuk mengetahui apakah residual sudah memenuhi asumsi *white noise* dan berdistribusi normal, seperti pada Tabel 4.31 berikut.

Tabel 4. 31 *Diagnostic Checking Model ARIMA Pos Kedungadem dengan Differencing Musiman 36*

Model ARIMA	Uji White Noise		Uji Normalitas
	Lag	P-value	
ARIMA (0,0,1) (0,1,1) ³⁶	6	0,0127	0,0220
	12	0,0886	
	18	0,1425	
	24	0,2832	
	30	0,4223	
	36	0,5522	
	42	0,7332	
	48	0,7375	
ARIMA (0,0,2) (0,1,1) ³⁶	6	0,3648	<0,0100
	12	0,6117	
	18	0,6270	
	24	0,7602	
	30	0,8306	
	36	0,8590	
	42	0,9352	
	48	0,9404	
ARIMA (1,0,1)(0,1,1) ³⁶	6	0,8194	<0,0100
	12	0,8435	
	18	0,7524	
	24	0,8436	
	30	0,8877	
	36	0,9086	
	42	0,9592	
	48	0,9556	

Tabel 4. 31 *Diagnostic Checking Model ARIMA Pos Kedungadem dengan Differencing Musiman 36 (Lanjutan)*

Model ARIMA	Uji <i>White Noise</i>		Uji Normalitas
	Lag	P-value	
ARIMA (1,0,[1,52])(0,1,1) ³⁶	6	0,5352	<0,0100
	12	0,7197	
	18	0,6254	
	24	0,7441	
	30	0,8241	
	36	0,8632	
	42	0,9282	
	48	0,9266	

Dari Tabel 4.31 didapatkan kesimpulan bahwa dari 4 model dugaan ARIMA terdapat 3 model yang memenuhi asumsi residual yang *white noise*. Hal ini ditunjukkan oleh nilai *P-value* yang lebih dari α ($\alpha=0,05$) sehingga gagal tolak H_0 . Selanjutnya dari 3 model dugaan tersebut akan dilakukan pengujian residual berdistribusi normal. Pada Tabel 4.31 juga menunjukkan *P-value* uji *Kolmogorovsmirnov* pada kolom uji normalitas. Dari pengujian normalitas residual didapatkan hasil yaitu dari 3 model ARIMA yang sebelumnya sudah memenuhi asumsi *white noise*, tidak memenuhi asumsi normalitas residual yang ditunjukkan oleh nilai *P-value* yang kurang dari α ($\alpha=0,05$), sehingga tolak H_0 . Dikarenakan model ARIMA tidak dapat memenuhi asumsi normalitas residual, akan dilakukan pendeteksian *outlier* model ARIMA. Hasil pendeteksian *outlier* terdapat pada Tabel 4.32.

Tabel 4. 32 *Deteksi Outlier Model ARIMA Pos Kedungadem dengan Differencing Musiman 36*

Model ARIMA	Pengamatan ke-	Jenis <i>Outlier</i>
ARIMA (0,0,2)(0,1,1) ³⁶	316	Additive
	355	Additive
	466	Additive
	24	Additive
	52	Additive

Tabel 4. 32 Deteksi *Outlier* Model ARIMA Pos Kedungadem dengan *Differencing* Musiman 36 (Lanjutan)

Model ARIMA	Pengamatan ke-	Jenis <i>Outlier</i>
ARIMA (1,0,1)(0,1,1) ³⁶	466	Additive
	316	Additive
	355	Additive
	265	Additive
	359	Additive
ARIMA (1,0,[1,52])(0,1,1) ³⁶	466	Additive
	355	Additive
	265	Additive
	316	Additive
	24	Additive

Berdasarkan Tabel 4.32 didapatkan hasil deteksi *outlier* pada tiga model dugaan ARIMA Pos Kedungadem dengan *differencing* musiman 36. Setelah mendapatkan pengamatan yang *outlier*, selanjutnya data *outlier* akan dimasukkan ke dalam model ARIMA sementara dan kemudian dilakukan estimasi parameter serta pengujian asumsi. Hasil pemodelan ARIMA dengan *outlier* terdapat pada Lampiran .

Pada Tabel 4.33 menunjukkan hasil pemodelan dengan penambahan *outlier* tidak memberikan hasil residual yang berdistribusi normal. Gejala ini tidak bisa diatasi dengan deteksi *outlier* dikarenakan histogram dari residual cenderung lebih runcing (leptokurtik) serta nilai kurtosisnya bernilai positif yang menunjukkan residual model ARIMA tidak berdistribusi normal.

Tabel 4. 33 Hasi Uji Normalitas Residual Model ARIMA *Outlier* Pos Kedungadem dengan *Differencing* Musiman 36

No	Model	<i>P-value</i>
1	ARIMA (0,0,2)(0,1,1) ³⁶	<0,0100
2	ARIMA (1,0,1)(0,1,1) ³⁶	<0,0100
3	ARIMA (1,0,[1,52])(0,1,1) ³⁶	<0,0100

Selanjutnya akan dilakukan perhitungan kriteria kebaikan model ARIMA berdasarkan pada RMSE dan SMAPE untuk memilih model ARIMA terbaik. Perhitungan kriteria kebaikan model terdapat pada Tabel 4.34 sebagai berikut.

Tabel 4. 34 Kriteria Kebaikan Model ARIMA Pos Kedungadem dengan Differencing Musiman 36

No	Model ARIMA	<i>In sample</i>	<i>Out sample</i>	
		RMSE	RMSE	SMAPE
1	ARIMA (0,0,2)(0,1,1) ³⁶	58,0898	37,3650	45,5300
2	ARIMA (1,0,1)(0,1,1) ³⁶	57,9267	36,8041	45,6412
3	ARIMA (1,0,[1,52])(0,1,1) ³⁶	57,7702	33,9779	43,5077

Dalam menentukan model ARIMA terbaik dapat menggunakan kriteria *in sample* dan kriteria *out sample*. Tabel 4.34 diketahui bahwa model ARIMA (1,0,[1,52])(0,1,1)³⁶ adalah model ARIMA terbaik yang ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil. Berikut adalah penguraian model ARIMA yang terdapat pada Tabel 4.34

ARIMA (0,0,2)(0,1,1)³⁶

$$Y_t^*(1 - B^{36}) = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{36} a_{t-36}$$

$$Y_t^* - Y_{t-36}^* = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{36} a_{t-36}$$

$$Y_t^* = Y_{t-36}^* + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_{36} a_{t-36}$$

$$Y_t^* = Y_{t-36}^* + a_t + 0,1441 a_{t-1} + 0,1210 a_{t-2} - 0,7317 a_{t-36}$$

ARIMA (1,0,1)(0,1,1)³⁶

$$Y_t^*(1 - \phi_1 B)(1 - B^{36}) = a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36}$$

$$Y_t^* - \phi_1 Y_{t-1}^* - Y_{t-36}^* + \phi_1 Y_{t-37}^* = a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36}$$

$$Y_t^* = \phi_1 Y_{t-1}^* + Y_{t-36}^* - \phi_1 Y_{t-37}^* + a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36}$$

$$Y_t^* = 0,6870 Y_{t-1}^* + Y_{t-36}^* - 0,6870 Y_{t-37}^* + a_t - 0,5319 a_{t-1} - 0,7296 a_{t-36}$$

ARIMA (1,0,[1,52])(0,1,1)³⁶

$$Y_t^*(1 - \phi_1 B)(1 - B^{36}) = a_t - \theta_1 a_{t-1} - \theta_{52} a_{t-52} - \theta_{36} a_{t-36}$$

$$\begin{aligned}
Y_t^* - \phi_1 Y_{t-1}^* - Y_{t-36}^* + \phi_1 Y_{t-37}^* \\
&= a_t - \theta_1 a_{t-1} - \theta_{52} a_{t-52} - \theta_{36} a_{t-36} \\
Y_t^* &= \phi_1 Y_{t-1}^* + Y_{t-36}^* - \phi_1 Y_{t-37}^* + a_t - \theta_1 a_{t-1} - \theta_{52} a_{t-52} \\
&\quad - \theta_{36} a_{t-36} \\
Y_t^* &= 0,6886 Y_{t-1}^* + Y_{t-36}^* - 0,6886 Y_{t-37}^* + a_t - 0,5322 a_{t-1} \\
&\quad + 0,0903 a_{t-52} - 0,7259 a_{t-36}
\end{aligned}$$

dimana Y_t^* adalah data curah hujan hasil transformasi

4.3.2 Peramalan Curah Hujan Menggunakan SVR

Tahap awal sebelum masuk pada metode *Support Vector Regression* (SVR) yakni dilakukan pengujian nonlinieritas pada data curah hujan di Pos Kedungadem. Pengujian nonlinieritas digunakan untuk memilih fungsi kernel pada metode SVR. Uji nonlinieritas akan menggunakan uji Terasvirta. Berdasarkan uji Terasvirta didapatkan P -value sebesar $2,931 \cdot 10^{-11}$. Pada uji Terasvirta ini dilakukan dalam beberapa *lag* yaitu *lag 1*, *lag 2*, *lag 3*, *lag 4*, *lag 5*, *lag 10* dan *lag 13* dari hasil pengujian tersebut didapatkan hasil P -value kurang dari α ($\alpha=5\%$), sehingga data curah hujan dasarian Pos Kedungadem memiliki pola non linier. Hal ini menunjukkan data curah hujan dasarian di Pos Kedungadem bersifat nonlinier, oleh karena itu dalam melakukan peramalan dengan SVR akan digunakan fungsi kernel *Radial Basis Function* (RBF).

Pada peramalan dengan SVR akan menggunakan *lag-lag input* yang didapatkan dari penguraian model ARIMA yang sudah memiliki parameter yang signifikan. Berdasarkan pada penguraian model ARIMA pada analisis sebelumnya dan pada lampiran 31 diperoleh beberapa *lag* Y_t . Berikut hasil penguraian beberapa model ARIMA yang sudah memiliki parameter yang signifikan beserta *lag input* Y_t .

Tabel 4. 35 *Input Lag SVR Pos Kedungadem*

No	Model ARIMA	Input Lag (Y_t)
1	ARIMA (2,0,0)	Y_{t-1}, Y_{t-2}
2	ARIMA ([1,2,17],0,0)	$Y_{t-1}, Y_{t-2}, Y_{t-17}$
3	ARIMA([8,10],0,0)	Y_{t-8}, Y_{t-10}
4	ARIMA([1,8],0,0)	Y_{t-1}, Y_{t-8}

Tabel 4. 35 *Input Lag* SVR Pos Kedugadem (Lanjutan)

No	Model ARIMA	<i>Input Lag</i> (Y_t)
5	ARIMA (2,0,0)(2,1,0) ¹⁸	$Y_{t-1}, Y_{t-2}, Y_{t-18}, Y_{t-19},$ $Y_{t-20}, Y_{t-36}, Y_{t-37}, Y_{t-38},$ $Y_{t-54}, Y_{t-55}, Y_{t-56}$
6	ARIMA ([1,2,16],0,0) (3,1,0) ¹⁸	$Y_{t-1}, Y_{t-2}, Y_{t-16}, Y_{t-18},$ $Y_{t-19}, Y_{t-20}, Y_{t-34}, Y_{t-36},$ $Y_{t-37}, Y_{t-38}, Y_{t-52}, Y_{t-54},$ $Y_{t-55}, Y_{t-56}, Y_{t-70}, Y_{t-72},$ $Y_{t-73}, Y_{t-74}, Y_{t-88}$
7	ARIMA ([1,16,37],0,0)(3,1,0) ¹⁸	$Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19},$ $Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52},$ $Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72},$ $Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109}$
8	ARIMA (0,0,1)(0,1,1) ³⁶	Y_{t-36}
9	ARIMA (0,0,2) (0,1,1) ³⁶	Y_{t-36}
10	ARIMA (1,0,1) (0,1,1) ³⁶	$Y_{t-1}, Y_{t-36}, Y_{t-37}$
11	ARIMA (1,0,[1,52]) (0,1,0) ³⁶	$Y_{t-1}, Y_{t-36}, Y_{t-37}$

Tabel 4.26, Tabel 4.28 dan Tabel 4.30 menunjukkan terdapat 11 model ARIMA Pos Kedugadem yang telah memiliki parameter signifikan. Tabel 4.35 menunjukkan *input lag* yang didapatkan dari penguraian model ARIMA. Berdasarkan *input lag* tersebut, maka *lag* tersebut akan menjadi *input* (X) pada peramalan dengan SVR. Setelah mendapatkan *input*, maka tahap selanjutnya adalah menentukan parameter model SVR.

Metode SVR dengan menggunakan fungsi kernel RBF terdapat 3 parameter yaitu C (Cost), γ dan ε . Namun, untuk mendapatkan hasil yang optimal parameter-parameter tersebut perlu diatur nilainya. Untuk mengatur nilai parameter tersebut, pada penelitian ini akan menggunakan variasi *trial and error*. Untuk memperoleh hasil peramalan yang baik, akan dikombinasikan beberapa pilihan nilai *range* parameter. Untuk mempermudah dalam pemilihan parameter maka terlebih dahulu akan dicari *range* optimal parameter ε yang optimal, dimana untuk parameter C dan γ ditetapkan *range* nya. Pada percobaan mencari parameter ε yang optimal, akan ditetapkan nilai untuk parameter C adalah 10^{-2} -10 dan

untuk γ adalah 0,5-1,5. Sebagai percobaan akan digunakan model input $f(Y_{t-1}, Y_{t-36}, Y_{t-37})$ untuk mendapatkan parameter yang optimal. Hasil percobaan nilai *range* parameter ε terdapat pada Tabel 4.36.

Tabel 4. 36 Hasil Percobaan *Range* Parameter ε Pos Kedungadem

Percobaan ke-	<i>Range</i> Parameter ε	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	0,0001-0,001	50,7137	47,2577	55,6624
2	0,001-0,01	50,6990	47,1725	55,5633
3	0,01-0,1	50,6537	47,0013	55,1607
4	0,1-0,5	50,2161	45,9755	54,2044
5	0,5-1	50,8015	45,5021	50,0515

Berdasarkan Tabel 4.36 didapatkan hasil *range* parameter ε yang optimal ketika nilai C dan γ tetap terdapat pada *range* 0,5-1, hal ini ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil dibandingkan pada *range* yang lain. Selanjutnya dengan cara yang sama akan dilakukan percobaan untuk mendapatkan *range* parameter γ yang optimal, dimana *range* untuk C adalah 10^{-2} -10 dan untuk parameter ε adalah 0,5-1. Berikut adalah hasil percobaan beberapa *range* parameter γ .

Tabel 4. 37 Hasil Percobaan *Range* Parameter γ Pos Kedungadem

Percobaan ke-	<i>Range</i> Parameter γ	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	0,05-0,1	52,8138	41,5980	48,8656
2	0,1-0,5	52,9255	41,6105	48,7002
3	0,5-1	50,8015	45,5021	50,0515
4	0,6-1,5	50,4405	45,2496	49,2722
5	1-1,5	49,3253	45,5696	49,2096

Tabel 4.37 menunjukkan bahwa *range* parameter γ yang optimal ketika nilai C dan ε tetap terdapat pada *range* 0,05-0,1 yang ditunjukkan oleh nilai RMSE dan SMAPE yang cukup kecil dibandingkan pada *range* yang lain. Langkah selanjutnya akan dilakukan percobaan untuk mendapatkan *range* parameter C yang optimal, dimana *range* untuk ε dan γ didapatkan dari percobaan se-

belumnya. Hasil beberapa percobaan *range* parameter C terdapat pada Tabel 4.38.

Tabel 4. 38 Hasil Percobaan *Range* Parameter C Pos Kedungadem

Percobaan	<i>Range</i>	RMSE	RMSE	SMAPE
ke-	Parameter C	<i>in sample</i>	<i>out sample</i>	<i>out sample</i>
1	0,01-10	52,7867	41,2816	48,5683
2	1-10	52,7831	41,2792	48,5687
3	5-50	52,8109	41,5964	48,8676
4	10-100	52,7466	41,4202	48,7239
5	100-150	52,1493	42,6596	49,1352

Berdasarkan Tabel 4.38 diperoleh informasi bahwa *range* parameter C yang optimal berda pada *range* 0,01-10. Dari beberapa percobaan yang telah dilakukan didapatkan hasil yakni parameter C yang optimal pada *range* 0,01-10, parameter ε pada *range* 0,5-1 dan parameter γ pada *range* 0,05-0,1. Selanjutnya setelah didapatkan *range* parameter yang optimal, maka *range* tersebut akan digunakan untuk memodelkan ARIMA yang telah memiliki parameter yang signifikan dengan metode SVR. Berikut merupakan hasil perhitungan kebaikan model dengan menggunakan metode SVR.

Tabel 4. 39 Kriteria Kebaikan Model SVR Pos Kedungadem

Model	RMSE	RMSE	SMAPE
	<i>in sample</i>	<i>out sample</i>	<i>out sample</i>
$f(Y_{t-1}, Y_{t-2})$	53,4628	36,2386	41,2593
$f(Y_{t-1}, Y_{t-2}, Y_{t-17})$	51,4148	36,8883	44,2713
$f(Y_{t-8}, Y_{t-10})$	61,3685	38,5177	40,5517
$f(Y_{t-1}, Y_{t-8})$	55,3269	38,7053	45,0523
$f(Y_{t-1}, Y_{t-2}, Y_{t-18},$ $Y_{t-19}, Y_{t-20}, Y_{t-36},$ $Y_{t-37}, Y_{t-38}, Y_{t-54}, Y_{t-55},$ $Y_{t-56})$	48,0288	36,3635	40,1549
$f(Y_{t-36})$	57,1360	32,1855	39,9169
$f(Y_{t-36})$	57,1360	32,1855	39,9169
$f(Y_{t-1}, Y_{t-36}, Y_{t-37})$	52,7867	41,2816	48,5682

Tabel 4. 39 Kriteria Keباikan Model SVR Pos Kedungadem (Lanjutan)

Model Input	RMSE in sample	RMSE out sample	SMAPE out sample
$f(Y_{t-1}, Y_{t-36}, Y_{t-37})$	52,7867	41,2816	48,5682
$f(Y_{t-1}, Y_{t-2}, Y_{t-16},$ $Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-34},$ $Y_{t-36}, Y_{t-37}, Y_{t-38},$ $Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-56}$ $, Y_{t-70}, Y_{t-72}, Y_{t-73},$ $Y_{t-74}, Y_{t-88})$	42,8774	44,0174	44,8926
$f(Y_{t-1}, Y_{t-16}, Y_{t-18},$ $Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}$ $, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}$ $, Y_{t-72}, Y_{t-73}, Y_{t-88},$ $Y_{t-91}, Y_{t-109})$	45,1607	36,5999	39,3409

Dari Tabel 4.39 didapatkan hasil yaitu model *input* $f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$ adalah model terbaik yang didapatkan dari metode SVR pada Pos Kedungadem. Hal ini ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil dari model yang lain. Model peramalan SVR terbaik berdasarkan data *out sample* adalah

$$f(x_i) = (\mathbf{w} \Phi(x_i) + b) = \sum_{i,j=1}^n \alpha_i K(x_i, x_j) + b$$

$$f(x_i) = (\mathbf{w} \Phi(x_i) + b) = \sum_{i,j=1}^{537} \alpha_i \exp(-0,05 \|x_i - x_j\|^2) + b$$

dengan $x_j = [x_{j1}, x_{j2}, \dots, x_{j16}]$ dan $x_i = [x_{i1}, x_{i2}, \dots, x_{i16}]$, berikut adalah persamaan model SVR terbaik yang dihasilkan berdasarkan Tabel 4.39.

$$f(x_i) = \sum_{i,j=1}^{467} \alpha_i \exp(-0,05 \|[x_{i1}, x_{i2}, \dots, x_{i16}] - [x_{j1}, x_{j2}, \dots, x_{j16}]\|^2) + b$$

Model SVR tersebut memiliki 467 data *out sample*, dimana parameter $C=1,01$, $\gamma= 0,05$ dan $\varepsilon= 0,5$. *Support Vector* yang dihasilkan sebanyak 206, artinya terdapat 206 data yang taksiran parameter α_i tidak bernilai nol.

4.3.3 Peramalan Curah Hujan Menggunakan GA-SVR

Setelah sebelumnya dilakukan peramalan SVR, maka perlu dilakukan optimasi parameter SVR agar menghasilkan peramalan yang lebih baik. Salah satu algoritma yang dapat digunakan adalah *Genetic Algorithm*. Pada *Genetic Algorithm* akan menggunakan 5 model SVR terbaik yang terbentuk pada langkah sebelumnya, kemudian model tersebut akan dioptimalkan parameternya menggunakan *Genetic Algorithm*. Langkah awal yang dilakukan adalah menginisialisasi kromosom sebanyak 100, nilai inisial kromosom didapatkan dari model SVR pada langkah sebelumnya, dimana setiap kromosom memiliki tiga gen. Tiga gen ini adalah parameter dari model SVR dengan menggunakan fungsi kernel RBF yaitu $Cost (C)$, γ dan ε . Misalkan digunakan model *input* SVR $f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$, dimana nilai nilai parameter yang didapatkan dari SVR adalah $C=1,01$, $\gamma=0,05$ dan $\varepsilon=0,5$. Berikut adalah ilustrasi kromosom dengan tiga gen.

Parameter	$Cost$	γ	ε
Kromosom	1,01	0,05	0,5

Gambar 4. 26 Ilustrasi Kromosom Parameter SVR dengan Tiga Gen Untuk Pos Kedungadem

Salah satu contoh kromosom yang dibentuk dari model SVR dengan variasi *trial and error* ditunjukkan oleh Gambar 4.26. Nilai parameter yang sebelumnya didapatkan dari SVR akan digunakan sebagai nilai inisial pada GA-SVR. Kromosom yang terbentuk menjalani proses seleksi *roulette wheel*, *crossover*, mutasi dan *elitism* hingga diperoleh hasil yang konvergen dengan nilai RMSE dan SMAPE yang lebih kecil.

Nilai *fitness* yang digunakan pada penelitian ini adalah nilai kesalahan dalam peramalan yaitu RMSE. Nilai *fitness* merupakan acuan dalam tahapan *Genetic Algorithm* seperti melakukan proses seleksi, pindah silang (*crossover*), mutasi dan *elitism* sehingga didapatkan nilai *fitness* terbaik. Fungsi objektif yang ingin dicapai

pada penelitian ini adalah meminimumkan nilai RMSE pada model SVR. Sebelum melewati tahapan dalam *Genetic Algorithm* maka akan dihitung terlebih dahulu nilai *fitness* berdasarkan kromosom-kromosom yang terbentuk. Pada ilustrasi akan dicontohkan pada model optimasi pada model $f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$. Ilustrasi nilai *fitness* masing-masing kromosom dapat dilihat pada Tabel 4.40.

Tabel 4. 40 Ilustrasi Nilai *Fitness* Setiap Kromosom Pos Kedungadem

Kromosom ke-	Kromosom			Nilai <i>Fitness</i>
	C	γ	ϵ	
1	1,0989	0,0588	0,4143	43,2665
2	1,0546	0,0523	0,4235	44,2054
:	:	:	:	:
100	1,0234	0,0543	0,4115	44,0164

Tabel 4.40 menunjukkan nilai *fitness* yang terdapat pada setiap kromosom memiliki nilai yang berbeda-beda. Nilai *fitness* tersebut akan menjadi acuan dalam melakukan proses GA selanjutnya hingga diperoleh nilai *fitness* yang paling minimum. Tahap selanjutnya adalah seleksi, dimana terdapat beberapa metode seleksi yang dapat digunakan. Penelitian ini akan menggunakan metode seleksi *roulette wheel*. Seleksi *roulette wheel* merupakan salah satu cara dalam penentuan orang tua kromosom yang akan melakukan *crossover* atau pindah silang. Orang tua yang terpilih dalam proses *roulette wheel* melibatkan nilai *fitness* relatif dan *fitness* kumulatif. Tabel 4.41 menunjukkan ilustrasi nilai *fitness* relatif dan *fitness* kumulatif.

Setelah didapatkan nilai *fitness* relatif dan *fitness* kumulatif, langkah selanjutnya adalah membangkitkan bilangan *random* [0,1] sebanyak jumlah kromosom yaitu 100. Bilangan *random* ini digunakan sebagai pembanding untuk setiap individu apakah individu tersebut layak untuk menjadi calon orang tua. Ilustrasi per-

bandingan bilangan *random* dengan nilai *fitness* kumulatif yang akan digunakan pada tahap seleksi terdapat pada Tabel 4.42.

Tabel 4. 41 Ilustrasi Nilai *Fitness* Relatif dan *Fitness* Kumulatif Pos Kedungadem

Kromosom ke-	<i>Fitness</i>	<i>Fitness Relatif</i>	<i>Fitness Kumulatif</i>
1	43,2665	0,0097	0,0097
2	44,2054	0,0096	0,0195
:	:	:	:
100	44,0164	0,0094	1

Tabel 4. 42 Ilustrasi Nilai Perbandingan *Fitness* Kumulatif dan Bilangan Random Pos Kedungadem

Kromosom ke-	<i>Fitness Kumulatif</i>	Bilangan Random
1	0,0097	0,0165
2	0,0195	0,9986
:	:	:
100	1	0,9765

Berdasarkan Tabel 4.42 diperoleh nilai bilangan random kromosom pertama adalah 0,0165. Nilai ini kurang dari kromosom ke-2, sehingga kromosom ke-2 terpilih sebagai calon orang tua. Pada kromosom kedua memiliki bilangan random kurang dari nilai *fitness* kumulatif pada kromosom ke-100, sehingga kromosom ke-100 terpilih menjadi calon orang tua pada proses selanjutnya.

Setelah proses seleksi adalah proses *crossover* atau pindah silang. *Crossover* melibatkan dua kromosom orang tua yang akan membentuk dua kromosom anak. *Crossover* dilakukan jika bilangan *random* lebih kecil dari probabilitas *crossover* ($P_c = 0,8$). Penelitian ini akan menggunakan tipe *crossover local arithmetic*. Tipe *crossover* ini banyak digunakan untuk kasus yang menggunakan nilai bilangan *real* (*real values encoding*) (Eiben, 2003). Berikut algoritma perhitungan *local arithmetic crossover*.

$$C_i^{gen+1} = \alpha C_i^{gen} + (1 - \alpha) C_j^{gen}$$

$$C_j^{gen+1} = \alpha C_j^{gen} + (1 - \alpha) C_i^{gen}$$

dimana C^{gen} adalah kromosom orang tua, C^{gen+1} adalah kromosom anak hasil *crossover* dan α adalah bobot, dimana nilai α diantara *range* 0 dan 1. Ilustrasi proses pindah silang kromosom orangtua satu dan kromosom orangtua dua yang akan menghasilkan anak 1 dan anak 2 terdapat pada Gambar 4.27.

Orang tua 1	1,0023	0,0488	0,4562
Orang tua 2	0,9987	0,0498	0,4765
Anak 1	1,00016	0,049394	0,468266
Anak 2	1,00084	0,049206	0,464434

Gambar 4. 27 Ilustrasi Kromosom yang Melakukan *Crossover* Pos Kedungadem

Gambar 4.27 menunjukkan ilustrasi kromosom orangtua dan kromosom hasil *crossover*. Misalkan setelah bilangan random dibangkitkan diperoleh nilai $\alpha = 0,4056$, kemudian nilai α akan dimasukkan dalam perhitungan rumus sebelumnya, sehingga dari proses *crossover* akan menghasilkan anak 1 dan anak 2. Langkah selanjutnya adalah proses mutasi. Proses memilih gen yang akan mengalami mutasi dilakukan dengan membangkitkan bilangan *random* antara 0 sampai 1 sebanyak gen dalam kromosom. Jika bilangan *random* lebih kecil dari probabilitas mutasi, $P_m = 0,01$ maka gen tersebut mengalami mutasi. Gambar 4.28 menunjukkan ilustrasi proses mutasi.

Mutasi bertujuan untuk mencegah terjadinya *convergence premature*. Apabila bilangan random yang dibangkitkan lebih kecil dari probabilitas mutasi (P_m) maka gen tersebut akan dilakukan mutasi dengan cara mengganti gen tersebut dengan bilangan acak,

dimana probabilitas mutasi dalam penelitian ini adalah 0,01. Berdasarkan Gambar 4.28 ditunjukkan bahwa gen ke-3 adalah gen yang dilakukan mutasi, karena pada gen tersebut memiliki bilangan random yang lebih kecil dari P_m sebesar 0,01, sehingga gen dalam kromosom tersebut akan diganti dengan bilangan acak.

Sebelum dimutasi	Bilangan Random	0,5361	0,3432	0,0021
	Kromosom	1,0234	0,0543	0,4115
Setelah dimutasi	Kromosom	1,0234	0,0543	0,4317

Gambar 4. 28 Ilustrasi Proses Mutasi Pos Kedungadem

Elitism adalah proses terakhir yang digunakan untuk memperoleh nilai *fitness* terbaik. Proses *elitism* ini akan mempertahankan estimasi parameter yang menghasilkan nilai *fitness* terkecil untuk generasi selanjutnya. Jumlah kromosom terbaik yang dipertahankan untuk generasi selanjutnya adalah 5 kromosom. Berikut disajikan ilustrasi proses *elitism* pada generasi pertama dan generasi kedua pada Tabel 4.43 dan Tabel 4.44.

Tabel 4. 43 Ilustrasi Proses *Elitism* Generasi Pertama Pos Kedungadem

Kromosom ke-	Kromosom			Nilai <i>Fitness</i>
	C	γ	ε	
1	1,0145	0,0516	0,4356	44,5141
2	1,0321	0,0482	0,3987	44,4850
:	:	:	:	:
100	0,9992	0,0508	0,5234	45,3840

Kromosom yang bertahan ke generasi selanjutnya adalah kromosom yang menghasilkan *fitness* terkecil. Lima kromosom dengan nilai *fitness* terbaik akan bertahan pada generasi kedua. Tabel 4.44 menunjukkan ilustrasi *elitism* generasi kedua .

Proses *elitism* pada generasi pertama menunjukkan kromosom pertama dan kromosom kedua merupakan kromosom *elitism* dari generasi pertama. Proses selanjutnya generasi kedua juga akan me-

lakukan proses *elitism* hingga generasi-generasi selanjutnya sampai kondisi *fitness* telah konvergen. Tabel 4.45 menunjukkan hasil optimasi beberapa model SVR yang telah didapatkan pada langkah sebelumnya.

Tabel 4. 44 Ilustrasi Proses *Elitism* Generasi Kedua Pos Kedungadem

Kromosom ke-	Kromosom			Nilai <i>Fitness</i>
	C	γ	ϵ	
1	1,0145	0,0516	0,4356	44,5141
2	1,0321	0,0482	0,3987	44,4850
:	:	:	:	:
100	1,0023	0.0567	0,5562	45,3252

Tabel 4. 45 Hasil Optimasi Parameter Model SVR Pos Kedungadem Menggunakan GA

No	Model <i>Input</i>	RMSE	RMSE	SMAPE
		<i>in sample</i>	<i>out sample</i>	<i>out sample</i>
1	$f(Y_{t-1}, Y_{t-2})$	53,4059	36,2795	41,1911
2	$f(Y_{t-1}, Y_{t-2}, Y_{t-17})$	51,1191	36,3896	43,5163
3	$f(Y_{t-1}, Y_{t-2}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-54}, Y_{t-55}, Y_{t-56})$	47,7961	36,1657	40,4743
4	$f(Y_{t-1}, Y_{t-2}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-56}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-74}, Y_{t-88})$	42,8520	41,1462	44,3972
5	$f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$	42,9440	35,2800	39,1441

Berdasarkan Tabel 4.45 diketahui bahwa dari 5 model SVR terbaik, maka model *input* $f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$ adalah model terbaik yang didapatkan dari metode GA-SVR. Hal ini ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil dibandingkan dengan model GA-SVR yang lain.

4.3.4 Perbandingan Model Terbaik Pos Kedungadem

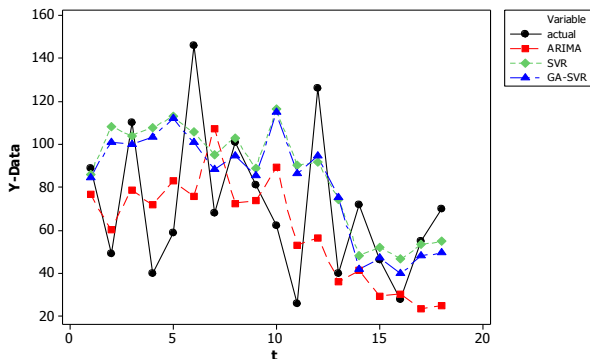
Langkah selanjutnya setelah didapatkan model terbaik dari ARIMA, SVR dan GA-SVR adalah melakukan perbandingan performansi model terbaik berdasarkan nilai RMSE *in sample*, RMSE *out sample* dan SMAPE *out sample*. Hasil perbandingan model terbaik ditunjukkan pada Tabel 4.46.

Tabel 4. 46 Hasil Perbandingan Metode Peramalan Pos Kedungadem

No	Metode	Model	RMSE <i>in sample</i>	RMSE <i>out sample</i>	SMAPE <i>out sample</i>
1	ARIMA	ARIMA (1,0,[1,52]) (0,1,1) ³⁶	57,7702	33,9779	43,5077
2	SVR	$f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$	45,1607	36,5999	39,3409
3	GA-SVR	$f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$	42,9440	35,2800	39,1441

Tabel 4.46 menunjukkan bahwa model terbaik untuk meramalkan curah hujan Pos Kedungadem adalah model $f(Y_{t-1},$

Y_{t-16} , Y_{t-18} , Y_{t-19} , Y_{t-34} , Y_{t-36} , Y_{t-37} , Y_{t-52} , Y_{t-54} , Y_{t-55} , Y_{t-70} , Y_{t-72} , Y_{t-73} , Y_{t-88} , Y_{t-91} , Y_{t-109}) yang didapatkan dari metode GA-SVR. Model ini memiliki ketepatan akurasi yang paling tinggi dibandingkan metode lainnya. Perbandingan secara visual pada data *out sample* terdapat pada Gambar 4.29. Tabel 4.47 menunjukkan *forecast* curah hujan 18 periode ke depan yakni pada bulan Juli 2016 sampai dengan Desember 2016 menggunakan metode GA-SVR pada Pos Kedungadem.



Gambar 4. 29 Perbandingan Visual pada Data *out sample* Pos Kedungadem

Tabel 4. 47 Hasil *Forecast* Curah Hujan 6 Periode Selanjutnya Pos Kedungadem

Dasarian	Bulan	Forecast
1	Juli	36,2111
2	Juli	22,6570
3	Juli	26.7335
1	Agustus	13,8875
2	Agustus	23,2287
3	Agustus	19,4592
1	September	12,0215
2	September	21,9582
3	September	17,3114

Tabel 4. 47 Hasil *Forecast* Curah Hujan 6 Periode Selanjutnya Pos Kedungadem (Lanjutan)

Dasarian	Bulan	<i>Forecast</i>
1	Oktober	13,2873
2	Oktober	20,3639
3	Oktober	29,5119
1	November	44,3715
2	November	67,2780
3	November	95,2588
1	Desember	74,9463
2	Desember	72,3197
3	Desember	55,2144

4.3.5 Kalender Tanam Padi Pos Kedungadem Periode Juli-Desember Tahun 2016

Hasil *forecast* curah hujan pada Pos Kedungadem bulan Juli hingga bulan Desember 2016, menunjukkan bahwa curah hujan selama 6 bulan tersebut antara 50-100 mm dan kurang dari 50 mm. Curah hujan ini termasuk pada kategori hujan sedang dan lebat (Suriadikusumah, 2007). Berikut adalah susunan kalender tanam padi untuk wilayah Kedungadem Kabupaten Bojonegoro bulan Juli 2016- Desember 2016.

Juli			Agustus			September			Oktober			November			Desember		
1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3



Musim tanam palawija

Musim tanam padi dengan bantuan pengairan

Gambar 4. 30 Kalender Tanam Padi Pos Kedungadem Periode Juli 2016 hingga Desember 2016

Gambar 4.30 menunjukkan kebutuhan air untuk penanaman padi sawah sepanjang bulan Juli 2016- Desember 2016 tidak dapat

terpenuhi, karena penanaman padi sawah membutuhkan air rata-rata 145 mm pada musim hujan (Oldeman, 1980), sehingga penanaman padi harus dibantu dengan pengairan dan tidak mengandalkan turunnya hujan. Pada bulan Juli hingga November petani dapat mengganti padi dengan palawija, dikarenakan palawija membutuhkan air yang lebih sedikit sekitar 50 mm pada musim kemarau (Oldeman, 1980). Jika persediaan cadangan air terbatas, maka petani dapat melakukan tanam palawija untuk mengantisipasi kekeringan atau kekurangan air. Jika dipaksakan untuk menanam padi, maka petani dan pemerintah harus memastikan tersedianya cadangan air dari embung atau waduk di daerah sekitar Pos Kedungadem.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis yang sudah dilakukan, maka didapatkan beberapa kesimpulan untuk menjawab tujuan yang telah ditetapkan sebagai berikut.

1. Karakteristik curah hujan di Pos Cawak dan Kedungadem menunjukkan bahwa pada bulan Januari, Februari, Maret, April, November dan Desember memiliki curah hujan yang tinggi. Sementara itu, pada bulan Juli, Agustus, September dan Oktober memiliki curah hujan yang rendah.
2. Pada peramalan menggunakan ARIMA menunjukkan bahwa model ARIMA terbaik untuk Pos Cawak adalah ARIMA $(0,0,1)(0,1,1)^{36}$, sedangkan model terbaik pada Pos Kedungadem adalah ARIMA $(1,0,[1,52]) (0,1,1)^{36}$.
3. Peramalan menggunakan SVR menghasilkan model terbaik untuk Pos Cawak dengan model $input f(Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-21}, Y_{t-37}, Y_{t-36}, Y_{t-38}, Y_{t-39})$ dan untuk Pos Kedungadem menggunakan input $f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$.
4. Setelah didapatkan beberapa model SVR, maka dilakukan optimasi dengan menggunakan *Genetic Algorithm* pada 5 model SVR terbaik. Ketika menggunakan GA-SVR diperoleh model terbaik yang sama dengan model SVR terbaik, namun akurasi peramalan lebih baik yang ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil.
5. Setelah dilakukan peramalan dengan menggunakan ARIMA, SVR dan GA-SVR, didapatkan hasil peramalan dengan GA-SVR menghasilkan peramalan dengan akurasi lebih baik.
6. Berdasarkan hasil *forecast* 6 bulan selanjutnya, dapat dibuat kalender tanam padi yang dapat digunakan pada tipe sawah tadah hujan. Hasil kalender tanam padi pada bulan Juli 2016-

Desember 2016 untuk kedua pos pengamatan menunjukkan kebutuhan air untuk penanaman padi sawah tidak dapat dipenuhi. Petani dapat mengganti dengan menanam palawija yang membutuhkan sedikit air. Jika tetap dilakukan tanam padi, maka petani dan pemerintah harus memastikan tersedianya cadangan air.

5.2 Saran

Penelitian yang telah dilakukan tentunya memiliki beberapa kekurangan, sehingga dari penelitian ini disarankan beberapa hal berikut untuk penelitian selanjutnya.

1. Dalam penelitian tugas akhir ini tidak memperhatikan adanya variabel lain yang mempengaruhi curah hujan sehingga disarankan untuk penelitian selanjutnya dapat digunakan metode yang bisa memperhatikan adanya pengaruh dari variabel lain.
2. Dapat digunakan kombinasi *range* parameter SVR yang lain untuk memperoleh hasil peramalan yang lebih baik.
3. Untuk penelitian selanjutnya diharapkan dapat dicoba menggunakan *input* dari data asli atau *input* yang lain tanpa melalui model ARIMA pada metode SVR.

DAFTAR PUSTAKA

- Alfredo, Jondri dan Rismala R. (2015). *Prediksi Harga Saham menggunakan Support Vector Regression dan Firefly Algorithm*. Bandung: Universitas Telkom.
- Badan Meteorologi, Klimatologi dan Geofisika (BMKG). (2016). *Tentang Meteorologi?*.
http://meteojuanda.info/index.php?option=com_content&view=article&id=36&Itemid=34. Diakses pada tanggal 12 November 2016.
- Badan Pusat Statistik. (2014). *Bojonegoro Dalam Angka Tahun 2014*. Bojonegoro: BPS.
- Badan Pusat Statistik. (2015). *Indikator Pertanian Jawa Timur Tahun 2015*. Surabaya: BPS.
- Badan Pusat Statistik. (2016). *Bojonegoro Dalam Angka Tahun 2016*. Bojonegoro : BPS.
- Bowerman, B., O'Connell, R. T., dan Koehler, A. B. (2004). *Forecasting, Time Series, and Regression: An Applied Approach*. Boston: Cengage Learning, Inc.
- Chen, K. Y. (2007). *Forecasting System Realibility Based on Support Vector Regression with Genetic Algorithm. Realibility Engineering and System Safety*.
- Cryer, J. D. dan Chan, K. (2008). *Time Series Analysis with Application in R*, 2nd Edition. New York: Springer.
- Daniel, W. W. (1989). *Statistika Nonparametrik Terapan*. Alih Bahasa: Alex Tri Kuncoro. Jakarta:PT Gramedia.
- Daurina L., (2015). *Kekeringan di Jatim, 248 Hektare Padi Gagal Panen*. <http://m.viva.co.id/ramadan2016/read/647348-kekeringan-di-jatim--248-hektare-padi-gagal-panen>. Diakses pada tanggal 5 September 2016.
- Dimiyati, H. (2014). *Support Vector Regression (SVR) untuk Menganalisa Pergerakan Indeks Harga Saham Gabungan (IHSG) Selama Pemilihan Umum 2014*. Thesis. Yogyakarta: Universitas Gadjah Mada.

- Eiben. (2003). *Introduction to Evolutionary Computation. Natural Computing Series*, Springer-Verlag Berlin Heiderberg.
- Gu, J., Zhu, M, dan Jiang L. (2011). Housing Price Forecasting Based on Genetic Algorithm and Support Vector Machine. *Expert System with Application*.
- Gujarati, D.N,. (2004). *Basic Econometrics, Fourth Edition*. New York: Mcgraw-Hill.
- Gunn, S.(1998). *Support Vector Machines for Classification. Technical Report*. Southampton: University of Southampton.
- Guo, X., Li, D. C. dan Zhang, A. (2012). Improved Support Vector Macchine Oil Price Forecast Model Based on Genetic Algorithm Optimization Parameters. *AASRI Conference on Computational Intelligence and Bioinformatics*, 525-530.
- Harafani, H. (2015). Optimasi Parameter pada Support Vector Machine Berbasis Algoritma Genetika untuk Estimasi Kebakaran Hutan. *Journal of Intelligent System*, Vol. 1, No. 2, Desember 2015.
- Huang, C.M., Lee, Y.-J., Lin, D. K., dan Huang, S.Y. (2007). Model Selection For Support Vector Machines Via Uniform Design. *Computational Statistics & Data Analysis*.
- Huda, A. M., Choiruddin, A., Budiarto, O dan Sutikno. (2012). Peramalan Data Curah Hujan Dengan Seasonal Autoregressive Integrated Moving Average (SARIMA) Dengan Deteksi Outlier Sebagai Upaya Optimalisasi Produksi Pertanian Di Kabupaten Mojokerto. *Seminar Nasional : Kedaulatan Pangan dan Energi*.
- Jia, Z.Y., Ma, W.J., Wang F. J dan Liu, W. (2011). Hybrid of Simulated Annealing and SVM for Hydraulic Valve Characteristics Prediction. *Expert Systems with Applications*.
- Lippi, M., Bertini, M dan Frasconi B. (2013). Short-Term Traffic Flow Forecasting: An Experimental Comparison of Time-

- Series Analysis and Supervised Learning. *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2.
- Makridakis S. dan Hibon, M. (2000). The M-3 Competition: Results, Conclusions and Implications. *International Journal of Forecasting*.
- Makridakis S., Wheelright S. C dan McGee V.E. (1999). *Metode dan Aplikasi Peramalan*. Edisi Kedua. Alih Bahasa: Untung Sus Andriyanto dan Abdul Basith. Jakarta: Erlangga.
- Msiza, I. S., Nelwamondo F. V. dan Marwala T., (2008). Water Demand Prediction using Artificial Neural Networks and Support Vector Regression. *Journal of Computers*, Vol. 3, No. 11, November 2008.
- Oldeman, L. R. (1980). *The Agroclimate Classification of Rice Growing Enviroment in Indonesia*.Philippines: IRRI.
- Patil, V.P. and Pawar D.D. (2015). The Optimal Crossover Or Mutation Rates In Genetic Algorithm : A Review. *International Journal of Applied Engineering and Technology*.
- Petrus, Soewono, C. N., Agung, A dan Sihana. (2007). Implementasi Metode Genetic Algorithm dan Simulated Annealing Dalam Optimasi Susunan Bahan Bakar Teras PWR Menggunakan Code Corebn. *Jurnal Tek.Reaktor Nuklir*, Vol. 11 No.3: Universitas Gadjah Mada.
- Putri, T. (2016). *Peramalan Return Saham Bank Central Asia Menggunakan Self Exciting Threshold Autoregressive-Genetic Algorithm*. Tugas Akhir. Surabaya: Institut Teknologi Sepuluh Nopember.
- Runtunuwu, E., Syahbuddin, H., Ramadhani, F., dan Kartiwa, B. (2013). *Pengembangan Sistem Informasi Kalender Tanam Terpadu*. Jakarta: Badan Penelitian dan Pengembangan Pertanian Kementrian Pertanian.
- Satyaning, P dan Irhamah. (2012). *Peramalan Beban Listrik Menggunakan Genetic Algorithm-Support Vector Machine (GA-SVM) di PT PLN (Persero) Sub Unit*

- Penyaluran dan Pusat Pengaturan Beban (P3B) Jawa Timur-Bali*. Tugas Akhir. Surabaya.: Institut Teknologi Sepuluh Nopember.
- Schaffer J.D, Caruana, R.A., Eshelman, L.J dan Das, R. (1989). A Study Of Control Parameters Affecting Online Performance Of Genetic Algorithms For Function Optimization. *Proceedings of the Third International Conference on Genetic Algorithms*.
- Suriadikusumah, A. (2007). *Analisis Curah Hujan Perhitungan & Penggunaannya*. Bandung: Rakayasa Sains.
- Suganyadevi, M. V, dan Babulal, C. K. (2014). Support Vector Regression Model For The Prediction Of Loadability Margin Of A Power System. *Applied Soft Computing Journal*.
- Waghoo, P dan Dhar S. (2013). Study of Optimal Mutation Factor for Genetic Algorithm. *International Journal of Application or Innovation in Engineering and Management*.
- Wei, W.,W.,S. (2006). *Time Series Analysis Univariate and Multivariate Method 2nd Editions*. New York: Addison Wesley Publishing Company, Inc.
- Wiyanti, DT dan Pulungan R. (2012). Peramalan Deret Waktu Menggunakan Model Fungso Basis Radial (RBF) dan Auto Regressive Integrated Moving Average (ARIMA). *Jurnal MIPA*. Semarang: Universitas Negeri Semarang.
- Yuliani, W., Irhamah dan Prastyo D D. (2014). *Penaksiran Parameter Model ARIMA dengan Menggunakan Genetic Algorithm*. Tugas Akhir. Surabaya: Institut Teknologi Sepuluh Nopember.
- Zhang, G P. (2003). Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. *Journal Neurocomputing*.
- Zhao, W., Tao, T., dan Zio, E. (2015). System Reliability Prediction by Support Vector Regression with Analytic

Selection and Genetic Algorithm Parameters Selection.
Applied Soft Computing.

(Halaman Ini Sengaja Dikosongkan)

LAMPIRAN

Lampiran 1

Data Curah Hujan Dasarian di Pos Cawak Bulan Januari 2000
hingga Bulan Juni 2016

No	Tahun	Bulan	Dasarian	Curah Hujan (mm)
1	2000	1	1	87
2	2000	1	2	164
3	2000	1	3	19
4	2000	2	1	125
5	2000	2	2	1
6	2000	2	3	6
7	2000	3	1	15
8	2000	3	2	85
9	2000	3	3	57
10	2000	4	1	55
11	2000	4	2	27
12	2000	4	3	69
13	2000	5	1	22
14	2000	5	2	0
15	2000	5	3	36
:	:	:	:	:
:	:	:	:	:
:	:	:	:	:
589	2016	5	1	10
590	2016	5	2	34
591	2016	5	3	159
592	2016	6	1	78
593	2016	6	2	110
594	2016	6	3	101

Lampiran 2

Data Curah Hujan Dasarian di Pos Kedungadem bulan Januari 2000 hingga bulan Juni 2016

No	Tahun	Bulan	Dasarian	Curah Hujan (mm)
1	2000	1	1	175
2	2000	1	2	98
3	2000	1	3	49
4	2000	2	1	69
5	2000	2	2	26
6	2000	2	3	57
7	2000	3	1	64
8	2000	3	2	74
9	2000	3	3	153
10	2000	4	1	22
11	2000	4	2	94
12	2000	4	3	57
13	2000	5	1	21
14	2000	5	2	5
15	2000	5	3	26
:	:	:	:	:
:	:	:	:	:
587	2016	4	2	11
588	2016	4	3	111
589	2016	5	1	25
590	2016	5	2	57
591	2016	5	3	31
592	2016	6	1	13
593	2016	6	2	40
594	2016	6	3	55

Lampiran 3

Syntax SAS ARIMA ([1,2,3,35],0,0) Pos Cawak

```
data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
0.086066
0.301511
:
0.316228
0.316228
0.316228
0.316228
0.102598
0.087370
0.235702
0.095783
0.316228
0.316228
;
proc arima data=datacawak2;
identify var=x;
estimate p=(1,2,3,35) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;
```

Lampiran 4*Syntax SAS ARIMA Outlier* ([1,2,3,35],0,0) Pos Cawak

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
:
0.316228
0.316228
;
data datacawak2;
set datacawak2;
if _n_=34 then ao1=1; else ao1=0;
if _n_=481 then ao2=1; else ao2=0;
if _n_=326 then ao3=1;else ao3=0;
if _n_=149 then ao4=1; else ao4=0;
if _n_=61 then ao5=1; else ao5=0;
if _n_=114 then ao6=1; else ao6=0;
if _n_=440 then ao7=1;else ao7=0;
if _n_=556 then ao8=1; else ao8=0;
if _n_=161 then ao9=1;else ao9=0;
if _n_>=575 then ls1=1; else ls2=0;
run;
proc arima data=datacawak2;
identify var=x crosscorr=(ao1 ao2 ao3 ao4 ao5 ao6 ao7 ao8 ao9
ls1);
estimate p=(1,2,3,35) input=(ao1 ao2 ao3 ao4 ao5 ao6 ao7 ao8
ao9 ls1) noconstant method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 5**Output ARIMA Outlier ([1,2,3,35],0,0) Pos Cawak**

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
AR1,1	0.43162	0.04104	10.52	<.0001	1	x	0
AR1,2	0.16664	0.04492	3.71	0.0002	2	x	0
AR1,3	0.16366	0.04037	4.05	<.0001	3	x	0
AR1,4	0.23617	0.02979	7.93	<.0001	35	x	0
NUM1	0.23006	0.06041	3.81	0.0002	0	ao1	0
NUM2	0.19447	0.06047	3.22	0.0014	0	ao2	0
NUM3	0.19295	0.06069	3.18	0.0016	0	ao3	0
NUM4	0.21819	0.06141	3.55	0.0004	0	ao4	0
NUM5	0.12690	0.06354	2.00	0.0463	0	ao5	0
NUM6	0.18260	0.06155	2.97	0.0031	0	ao6	0
NUM7	0.17064	0.06051	2.82	0.0050	0	ao7	0
NUM8	-0.17057	0.06214	-2.74	0.0062	0	ao8	0
NUM9	-0.17423	0.06055	-2.88	0.0042	0	ao9	0
NUM10	-0.01259	0.03898	-0.32	0.7468	0	ls1	0

The ARIMA Procedure

Autocorrelation Check of Residuals

To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	3.62	2	0.1640	-0.022	-0.040	-0.032	0.039	0.039	0.009
12	9.82	8	0.2781	0.064	-0.004	0.039	-0.007	0.019	0.068
18	23.53	14	0.0521	0.075	0.062	-0.097	-0.025	-0.049	0.036
24	24.87	20	0.2066	-0.004	-0.011	0.004	-0.025	-0.038	-0.000
30	25.38	26	0.4974	0.020	0.017	0.004	0.004	-0.010	0.003
36	35.82	32	0.2937	0.044	0.055	0.072	-0.034	-0.075	-0.005
42	43.06	38	0.2636	0.097	0.001	-0.024	0.033	0.008	-0.024
48	47.20	44	0.3432	0.012	-0.040	0.062	0.010	-0.008	-0.030

Tests for Normality

Test	--Statistic--	-----p Value-----
Shapiro-Wilk	W 0.992587	Pr < W 0.0059
Kolmogorov-Smirnov	D 0.053419	Pr > D <0.0100
Cramer-von Mises	W-Sq 0.384471	Pr > W-Sq <0.0050
Anderson-Darling	A-Sq 1.815844	Pr > A-Sq <0.0050

Lampiran 6*Syntax SAS ARIMA (1,0,1)(0,1,1)³⁶ Pos Cawak*

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
0.086066
0.301511
:
0.316228
0.316228
0.316228
0.316228
0.102598
0.087370
0.235702
0.095783
0.316228
0.316228
;
proc arima data=datacawak2;
identify var=x(36);
estimate p=(1) q=(1)(36) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 7*Syntax SAS ARIMA Outlier (1,0,1)(0,1,1)³⁶ Pos Cawak*

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
:
0.087370
0.235702
0.095783
0.316228
0.316228
;
data datacawak2;
set datacawak2;
if _n_=334 then ao1=1; else ao1=0;
if _n_=34 then ao2=1; else ao2=0;
if _n_=114 then ao3=1;else ao3=0;
if _n_=392 then ao4=1; else ao4=0;
if _n_>=575 then ls1=1;else ls1=0;
run;
proc arima data=datacawak2;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ao3(36) ao4(36)
ls1(36));
estimate p=(1) q=(1)(36) input=(ao1 ao2 ao3 ao4 ls1)
noconstant method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 8*Output ARIMA Outlier (1,0,1)(0,1,1)³⁶ Pos Cawak*

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
MA1,1	0.46779	0.14887	3.14	0.0018	1	x	0
MA2,1	0.74883	0.02978	25.14	<.0001	36	x	0
AR1,1	0.64892	0.12853	5.05	<.0001	1	x	0
NUM1	0.18382	0.06633	2.77	0.0058	0	ao1	0
NUM2	0.19818	0.04711	4.21	<.0001	0	ao2	0
NUM3	0.16756	0.06622	2.53	0.0117	0	ao3	0
NUM4	0.17806	0.06686	2.66	0.0080	0	ao4	0
NUM5	0.19087	0.05593	3.41	0.0007	0	ls1	0

Autocorrelation Check of Residuals

To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----						
6	1.07	3	0.7832	0.014	-0.038	0.014	-0.002	0.010	-0.007	
12	4.61	9	0.8668	0.041	-0.001	0.011	-0.027	0.059	0.019	
18	13.04	15	0.5995	0.075	0.029	-0.088	0.029	0.007	0.008	
24	15.49	21	0.7977	0.033	-0.006	-0.037	0.027	-0.034	0.001	
30	16.20	27	0.9491	-0.026	0.020	0.002	-0.012	-0.004	-0.001	
36	23.06	33	0.9013	-0.007	-0.022	0.044	-0.096	-0.010	0.008	
42	25.41	39	0.9543	0.046	0.001	-0.008	0.024	-0.001	-0.035	
48	29.63	45	0.9625	-0.046	-0.062	0.023	-0.006	-0.013	-0.021	

Tests for Normality

Test	--Statistic--		-----p Value-----	
Shapiro-Wilk	W	0.99344	Pr < W	0.0189
Kolmogorov-Smirnov	D	0.046587	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.203444	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.164495	Pr > A-Sq	<0.0050

Lampiran 9*Syntax SAS ARIMA (0,0,1)(0,1,1)³⁶ Pos Cawak*

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
0.086066
0.301511
:
0.316228
0.316228
0.316228
0.316228
0.102598
0.087370
0.235702
0.095783
0.316228
0.316228
;
proc arima data=datacawak2;
identify var=x(36);
estimate q=(1)(36) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 10*Syntax SAS ARIMA Outlier (0,0,1)(0,1,1)³⁶ Pos Cawak*

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
0.086066
:
0.087370
0.235702
0.095783
0.316228
0.316228
;
data datacawak2;
set datacawak2;
if _n_=334 then ao1=1; else ao1=0;
if _n_=34 then ao2=1; else ao2=0;
if _n_=114 then ao3=1;else ao3=0;
if _n_=209 then ao4=1; else ao4=0;
if _n_>=575 then ls1=1;else ls1=0;
run;
proc arima data=datacawak2;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ao3(36) ao4(36)
ls1(36));
estimate q=(1)(36) input=(ao1 ao2 ao3 ao4 ls1) noconstant
method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```


Lampiran 11**Output ARIMA Outlier (0,0,1)(0,1,1)³⁶ Pos Cawak**

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
MA1,1	-0.19855	0.04261	-4.66	<.0001	1	x	0
MA2,1	0.74765	0.02987	25.03	<.0001	36	x	0
NUM1	0.20070	0.06668	3.01	0.0027	0	ao1	0
NUM2	0.20347	0.04730	4.30	<.0001	0	ao2	0
NUM3	0.17433	0.06655	2.62	0.0091	0	ao3	0
NUM4	-0.16925	0.06686	-2.53	0.0117	0	ao4	0
NUM5	0.19700	0.05669	3.47	0.0006	0	ls1	0

Autocorrelation Check of Residuals

To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	6.52	4	0.1638	0.013	0.076	0.064	0.028	0.033	0.003
12	10.07	10	0.4348	0.052	-0.002	0.020	-0.017	0.052	0.016
18	18.74	16	0.2822	0.067	0.057	-0.082	0.030	-0.002	-0.013
24	20.28	22	0.5654	0.032	-0.012	-0.022	0.004	-0.032	0.005
30	21.42	28	0.8075	-0.025	0.027	-0.010	-0.011	-0.016	-0.013
36	30.09	34	0.6600	-0.003	-0.035	0.063	-0.098	0.010	-0.006
42	32.47	40	0.7956	0.045	0.007	-0.001	0.016	-0.003	-0.041
48	36.89	46	0.8291	-0.044	-0.057	0.025	-0.034	-0.004	-0.020

Tests for Normality

Test	--Statistic--	-----p Value-----
Shapiro-Wilk	W 0.994755	Pr < W 0.0621
Kolmogorov-Smirnov	D 0.04656	Pr > D <0.0100
Cramer-von Mises	W-Sq 0.197194	Pr > W-Sq 0.0058
Anderson-Darling	A-Sq 1.048566	Pr > A-Sq 0.0094

Lampiran 12*Syntax SAS ARIMA (0,0,2)(0,1,1)³⁶ Pos Cawak*

```
data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
0.086066
0.301511
:
0.316228
0.316228
0.316228
0.316228
0.102598
0.087370
0.235702
0.095783
0.316228
0.316228
;
proc arima data=datacawak2;
identify var=x(36);
estimate q=(1,2)(36) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;
```

Lampiran 13*Syntax SAS ARIMA Outlier (0,0,2)(0,1,1)³⁶ Pos Cawak*

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
:
0.087370
0.235702
0.095783
0.316228
0.316228
;
data datacawak2;
set datacawak2;
if _n_=334 then ao1=1; else ao1=0;
if _n_=34 then ao2=1; else ao2=0;
if _n_=114 then ao3=1;else ao3=0;
if _n_=440 then ao4=1; else ao4=0;
if _n_>=575 then ls1=1;else ls1=0;
run;
proc arima data=datacawak2;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ao3(36) ao4(36)
ls1(36));
estimate q=(1,2)(36) input=(ao1 ao2 ao3 ao4 ls1) noconstant
method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 14*Output ARIMA Outlier (0,0,2)(0,1,1)³⁶ Pos Cawak*

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
MA1,1	-0.19329	0.04329	-4.47	<.0001	1	x	0
MA1,2	-0.07895	0.04336	-1.82	0.0692	2	x	0
MA2,1	0.74841	0.02992	25.01	<.0001	36	x	0
NUM1	0.19304	0.06657	2.90	0.0039	0	ao1	0
NUM2	0.19394	0.04716	4.11	<.0001	0	ao2	0
NUM3	0.17618	0.06646	2.65	0.0083	0	ao3	0
NUM4	0.18463	0.06738	2.74	0.0063	0	ao4	0
NUM5	0.19343	0.05644	3.43	0.0007	0	ls1	0

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	3.47	3	0.3240	0.005	0.013	0.068	0.020	0.032	0.012
12	7.05	9	0.6315	0.043	-0.003	0.024	-0.020	0.056	0.024
18	14.88	15	0.4602	0.068	0.048	-0.080	0.025	-0.010	0.003
24	16.74	21	0.7266	0.027	0.002	-0.034	0.015	-0.033	0.005
30	18.24	27	0.8962	-0.028	0.036	-0.012	-0.017	-0.006	-0.010
36	27.91	33	0.7185	-0.003	-0.026	0.063	-0.110	0.005	0.001
42	30.46	39	0.8342	0.046	-0.003	-0.001	0.024	-0.020	-0.036
48	33.73	45	0.8913	-0.037	-0.058	0.021	-0.013	-0.011	-0.012

Tests for Normality

Test	--Statistic--	-----p Value-----
Shapiro-Wilk	W 0.994252	Pr < W 0.0394
Kolmogorov-Smirnov	D 0.045665	Pr > D <0.0100
Cramer-von Mises	W-Sq 0.19235	Pr > W-Sq 0.0068
Anderson-Darling	A-Sq 1.04802	Pr > A-Sq 0.0094

Lampiran 15*Syntax SAS ARIMA ([1,36,37],0,0)(0,1,0)³⁶ Pos Cawak*

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
0.086066
0.301511
:
0.316228
0.316228
0.316228
0.316228
0.102598
0.087370
0.235702
0.095783
0.316228
0.316228
;
proc arima data=datacawak2;
identify var=x(36);
estimate p=(1,36,37) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 16

Syntax SAS ARIMA Outlier ([1,36,37],0,0)(0,1,0)³⁶ Pos Cawak

```

data datacawak2;
input x;
datalines;
0.101535
0.075810
0.185695
:

0.235702
0.095783
0.316228
0.316228
;
data datacawak2;
set datacawak2;
if _n_=114 then ao1=1; else ao1=0;
if _n_=172 then ao2=1; else ao2=0;
if _n_>=360 then ls1=1;else ls1=0;
if _n_>=575 then ls2=1;else ls2=0;
if _n_>=230 then ls3=1;else ls3=0;
run;
proc arima data=datacawak2;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ls1(36) ls2(36)
ls3(36));
estimate p=(1,36,37) input=(ao1 ao2 ls1 ls2 ls3) noconstant
method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 17*Output ARIMA Outlier ([1,36,37],0,0)(0,1,0)³⁶ Pos Cawak*

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
AR1,1	0.18314	0.04259	4.30	<.0001	1	x	0
AR1,2	-0.47011	0.03977	-11.82	<.0001	36	x	0
AR1,3	0.14770	0.04443	3.32	0.0009	37	x	0
NUM1	0.20333	0.06275	3.24	0.0013	0	ao1	0
NUM2	-0.18385	0.06276	-2.93	0.0035	0	ao2	0
NUM3	-0.05240	0.01484	-3.53	0.0005	0	ls1	0
NUM4	0.20582	0.06061	3.40	0.0007	0	ls2	0
NUM5	0.04889	0.01487	3.29	0.0011	0	ls3	0

Autocorrelation Check of Residuals

To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	1.48	3	0.6864	-0.004	0.016	0.040	0.017	0.009	-0.021
12	6.15	9	0.7245	0.062	0.022	0.023	-0.003	0.040	0.044
18	12.73	15	0.6231	0.065	0.049	-0.069	0.022	0.000	0.000
24	13.42	21	0.8932	-0.012	0.005	-0.009	-0.018	-0.004	0.025
30	14.91	27	0.9707	-0.003	0.028	-0.001	0.038	0.014	0.015
36	31.68	33	0.5330	0.012	-0.032	0.085	-0.092	0.010	-0.110
42	37.96	39	0.5171	-0.001	0.032	-0.041	0.082	0.007	-0.036
48	40.94	45	0.6446	-0.041	-0.033	0.046	-0.000	0.007	0.008

Tests for Normality

Test	--Statistic--		-----p Value-----	
Shapiro-Wilk	W	0.990554	Pr < W	0.0016
Kolmogorov-Smirnov	D	0.060113	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.512755	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	2.441992	Pr > A-Sq	<0.0050

Lampiran 18*Syntax SVR Model ARIMA (3,0,0)(1,1,0)*¹⁸ Pos Cawak

```

library(e1071)
library(tseries)
data=read.csv("D:/cawak5.csv",sep=",",header=T)
attach(data)
train=data[1:537,]
test=data[-c(1:537),]
tuneResult=tune(svm, Yt ~.,data=train,kernel="radial",ranges =
list(epsilon = seq(0.5,1), cost =seq(0.01,10),gamma=seq(1,1.5)))
tunedmodel=tuneResult$best.model;tunedmodel
forecast.in=predict(tunedmodel,train)
forecast.out=predict(tunedmodel,test)

forecast1=(forecast.in)
forecast2=(forecast.out)
testing=((test[,1]))
training=((train[,1]))
isi1=testing-forecast2
isi2=training-forecast1

rmse.in=sqrt(mean(isi2^2))
rmse.out=sqrt(mean(isi1^2))
smape.out=(sum(abs(testing-
forecast2)/((abs(testing)+abs(forecast2))/2))/length(testing))*1
00

write.table(forecast1,"D:/prediksisvrin.txt")
write.table(forecast2,"D:/prediksisvrout.txt")
rmse.in
rmse.out
smape.out

```


Lampiran 19

Hasil Penguraian Model ARIMA yang sudah memiliki parameter yang signifikan Pos Cawak

ARIMA (3,0,0)

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_3 Y_{t-3}^* &= a_t \\ Y_t^* &= a_t + \phi_1 Y_{t-1}^* + \phi_2 Y_{t-2}^* + \phi_3 Y_{t-3}^* \\ Y_t^* &= 0,4821 Y_{t-1}^* + 0,2643 Y_{t-2}^* + 0,2212 Y_{t-3}^* + a_t \end{aligned}$$

ARIMA ([8,10],0,0)

$$\begin{aligned} Y_t^* (1 - \phi_8 B^8 - \phi_{10} B^{10}) &= a_t \\ Y_t^* - \phi_8 Y_{t-8}^* - \phi_{10} Y_{t-10}^* &= a_t \\ Y_t^* &= a_t + \phi_8 Y_{t-8}^* + \phi_{10} Y_{t-10}^* \\ Y_t^* &= 0,5845 Y_{t-8}^* + 0,2889 Y_{t-10}^* + a_t \end{aligned}$$

ARIMA ([10,12,35],0,0)

$$\begin{aligned} Y_t^* (1 - \phi_{10} B^{10} - \phi_{12} B^{12} - \phi_{35} B^{35}) &= a_t \\ Y_t^* - \phi_{10} Y_{t-10}^* - \phi_{12} Y_{t-12}^* - \phi_{35} Y_{t-35}^* &= a_t \\ Y_t^* &= a_t + \phi_{10} Y_{t-10}^* + \phi_{12} Y_{t-12}^* + \phi_{35} Y_{t-35}^* \\ Y_t^* &= 0,2168 Y_{t-10}^* + 0,1308 Y_{t-12}^* + 0,6367 Y_{t-35}^* + a_t \end{aligned}$$

ARIMA ([1,2,3,35],0,0)

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_{35} B^{35}) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_3 Y_{t-3}^* - \phi_{35} Y_{t-35}^* &= a_t \\ Y_t^* &= a_t + \phi_1 Y_{t-1}^* + \phi_2 Y_{t-2}^* + \phi_3 Y_{t-3}^* + \phi_{35} Y_{t-35}^* \\ Y_t^* &= 0,4137 Y_{t-1}^* + 0,1890 Y_{t-2}^* + 0,1698 Y_{t-3}^* + 0,2200 Y_{t-35}^* + a_t \end{aligned}$$

ARIMA (3,0,0)(1,1,0)¹⁸

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) (1 - \Phi_{18} B^{18}) (1 - B^{18}) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_3 Y_{t-3}^* - Y_{t-18}^* + \phi_1 Y_{t-19}^* + \phi_2 Y_{t-20}^* \\ &\quad + \phi_3 Y_{t-21}^* - \Phi_{18} Y_{t-18}^* + \phi_1 \Phi_{18} Y_{t-19}^* \\ &\quad + \phi_2 \Phi_{18} Y_{t-20}^* + \phi_3 \Phi_{18} Y_{t-21}^* - \Phi_{18} Y_{t-36}^* \\ &\quad - \phi_1 \Phi_{18} Y_{t-37}^* - \phi_2 \Phi_{18} Y_{t-38}^* - \phi_3 \Phi_{18} Y_{t-39}^* = a_t \\ Y_t^* &= \phi_1 Y_{t-1}^* + \phi_2 Y_{t-2}^* + \phi_3 Y_{t-3}^* + Y_{t-18}^* - \phi_1 Y_{t-19}^* - \phi_2 Y_{t-20}^* \\ &\quad - \phi_3 Y_{t-21}^* + \Phi_{18} Y_{t-18}^* - \phi_1 \Phi_{18} Y_{t-19}^* \\ &\quad - \phi_2 \Phi_{18} Y_{t-20}^* - \phi_3 \Phi_{18} Y_{t-21}^* + \Phi_{18} Y_{t-36}^* \\ &\quad + \phi_1 \Phi_{18} Y_{t-37}^* + \phi_2 \Phi_{18} Y_{t-38}^* + \phi_3 \Phi_{18} Y_{t-39}^* + a_t \end{aligned}$$

$$\begin{aligned}
Y_t^* = & 0,2816 Y_{t-1}^* + 0,1885 Y_{t-2}^* + 0,1022 Y_{t-3}^* + Y_{t-18}^* \\
& - 0,2816 Y_{t-19}^* - 0,1885 Y_{t-20}^* - 0,1022 Y_{t-21}^* \\
& - 0,6479 Y_{t-18}^* + 0,1824 Y_{t-19}^* + 0,1221 Y_{t-20}^* \\
& + 0,0662 Y_{t-21}^* - 0,6479 Y_{t-36}^* - 0,1824 Y_{t-37}^* \\
& - 0,1221 Y_{t-38}^* - 0,0662 Y_{t-39}^* + a_t
\end{aligned}$$

ARIMA (3,0,0)(2,1,0)¹⁸

$$\begin{aligned}
Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) (1 - \Phi_{18} B^{18} - \Phi_{36} B^{36}) (1 - B^{18}) = a_t \\
Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_3 Y_{t-3}^* - \Phi_{18} Y_{t-18}^* + \phi_1 \Phi_{18} Y_{t-19}^* \\
+ \phi_2 \Phi_{18} Y_{t-20}^* + \phi_3 \Phi_{18} Y_{t-21}^* - \Phi_{36} Y_{t-36}^* \\
+ \phi_1 \Phi_{36} Y_{t-37}^* + \phi_2 \Phi_{36} Y_{t-38}^* + \phi_3 \Phi_{36} Y_{t-39}^* \\
- Y_{t-18}^* + \phi_1 Y_{t-19}^* + \phi_2 Y_{t-20}^* + \phi_3 Y_{t-21}^* \\
+ \Phi_{18} Y_{t-36}^* - \phi_1 \Phi_{18} Y_{t-37}^* - \phi_2 \Phi_{18} Y_{t-38}^* \\
- \phi_3 \Phi_{18} Y_{t-39}^* + \Phi_{36} Y_{t-54}^* - \phi_1 \Phi_{36} Y_{t-55}^* \\
- \phi_2 \Phi_{36} Y_{t-56}^* - \phi_3 \Phi_{36} Y_{t-57}^* = a_t
\end{aligned}$$

$$\begin{aligned}
Y_t^* = & 0,3272 Y_{t-1}^* + 0,2076 Y_{t-2}^* + 0,1143 Y_{t-3}^* - 0,7345 Y_{t-18}^* \\
& + 0,2403 Y_{t-19}^* + 0,1525 Y_{t-20}^* + 0,0840 Y_{t-21}^* \\
& - 0,1937 Y_{t-36}^* + 0,0634 Y_{t-37}^* + 0,0402 Y_{t-38}^* \\
& + 0,0221 Y_{t-39}^* + Y_{t-18}^* - 0,3272 Y_{t-19}^* \\
& - 0,2076 Y_{t-20}^* - 0,1143 Y_{t-21}^* + 0,7345 Y_{t-36}^* \\
& - 0,2403 Y_{t-37}^* - 0,1525 Y_{t-38}^* - 0,0840 Y_{t-39}^* \\
& + 0,1937 Y_{t-54}^* - 0,0634 Y_{t-55}^* - 0,0402 Y_{t-56}^* \\
& + 0,0221 Y_{t-57}^* + a_t
\end{aligned}$$

ARIMA ([1,2,3,15],0,0)(2,1,0)¹⁸

$$\begin{aligned}
Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_{15} B^{15}) (1 - \Phi_{18} B^{18} \\
- \Phi_{36} B^{36}) (1 - B^{18}) = a_t \\
Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_3 Y_{t-3}^* - \phi_{15} Y_{t-15}^* - \Phi_{18} Y_{t-18}^* \\
+ \phi_1 \Phi_{18} Y_{t-19}^* + \phi_2 \Phi_{18} Y_{t-20}^* + \phi_3 \Phi_{18} Y_{t-21}^* \\
+ \phi_{15} \Phi_{18} Y_{t-33}^* - \Phi_{36} Y_{t-36}^* + \phi_1 \Phi_{36} Y_{t-37}^* \\
+ \phi_2 \Phi_{36} Y_{t-38}^* + \phi_3 \Phi_{36} Y_{t-39}^* + \phi_{15} \Phi_{36} Y_{t-51}^* \\
- Y_{t-18}^* + \phi_1 Y_{t-19}^* + \phi_2 Y_{t-20}^* + \phi_3 Y_{t-21}^* \\
+ \phi_{15} Y_{t-33}^* + \Phi_{18} Y_{t-36}^* - \phi_1 \Phi_{18} Y_{t-37}^* \\
- \phi_2 \Phi_{18} Y_{t-38}^* - \phi_3 \Phi_{18} Y_{t-39}^* - \phi_{15} \Phi_{18} Y_{t-51}^* \\
+ \Phi_{36} Y_{t-54}^* - \phi_1 \Phi_{36} Y_{t-55}^* - \phi_2 \Phi_{36} Y_{t-56}^* \\
- \phi_3 \Phi_{36} Y_{t-57}^* - \phi_{15} \Phi_{36} Y_{t-69}^* = a_t
\end{aligned}$$

$$\begin{aligned}
Y_t^* = & 0,3250 Y_{t-1}^* + 0,1905 Y_{t-2}^* + 0,1278 Y_{t-3}^* - 0,2400 Y_{t-15}^* \\
& - 0,6508 Y_{t-18}^* + 0,2115 Y_{t-19}^* + 0,1240 Y_{t-20}^* \\
& + 0,0832 Y_{t-21}^* - 0,1562 Y_{t-33}^* - 0,2478 Y_{t-36}^* \\
& + 0,0805 Y_{t-37}^* + 0,0472 Y_{t-38}^* + 0,0317 Y_{t-39}^* \\
& - 0,0595 Y_{t-51}^* + Y_{t-18}^* - 0,3250 Y_{t-19}^* \\
& - 0,1905 Y_{t-20}^* - 0,1278 Y_{t-21}^* + 0,2400 Y_{t-33}^* \\
& + 0,6508 Y_{t-36}^* - 0,2115 Y_{t-37}^* - 0,1240 Y_{t-38}^* \\
& - 0,0832 Y_{t-39}^* + 0,1562 Y_{t-51}^* + 0,2478 Y_{t-54}^* \\
& - 0,0805 Y_{t-55}^* - 0,0472 Y_{t-56}^* - 0,0317 Y_{t-57}^* \\
& + 0,0595 Y_{t-69}^* + a_t
\end{aligned}$$

Lampiran 20

Syntax GA-SVR Model ARIMA Pos Cawak

```

library(e1071)
library(GA)
d=read.csv("d:/cawak9.csv",sep="," ,header=TRUE)
train_data = d[1:540,]
test_data = d[-c(1:540),]
# Fitness function (to be maximized)
# Parameter vector x is: (cost, gamma, epsilon)
fitnessFunc <- function(x) {
  # Retrieve the SVM parameters
  cost <- x[1]
  gamma <- x[2]
  epsilon <- x[3]
  # Train
  model <- svm(Yt ~ ., data = train_data, cost = cost, gamma = gamma, epsilon
= epsilon, type = "eps-regression", kernel = "radial")
  # Test
  rmse <- mean((predict(model, newdata = train_data) - train_data$Yt) ^ 2)
  return (-rmse)
}
# Range of the parameter values to be tested
# Parameters are: (cost, gamma, epsilon)
theta_min <- c(cost = 15.5, gamma = 0.989, epsilon = 0.4999)
theta_max <- c(cost = 16.5, gamma = 1.0199, epsilon = 0.51999)

# Run the genetic algorithm
results <- ga(type = "real-valued", fitness = fitnessFunc,
  names = names(theta_min),
  min = theta_min, max = theta_max,
  popSize = 100, maxiter = 100, prossover=0.8, pmuation=0.01, monitor=plot)
summary(results)
modpred=svm(Yt~.,data=train_data,cost=15.94504 ,gamma= 0.9926564
,epsilon=0.5178506)
forecast.in=predict(modpred,train_data)
forecast.out=predict(modpred,test_data)
forecast1=(forecast.in)
forecast2=(forecast.out)
zt=(test_data[,1])
xt=(train_data[,1])
isi1=xt-forecast1
isi2=zt-forecast2
rmse.in=sqrt(mean(isi1^2))
rmse.out=sqrt(mean(isi2^2))
smape.out=sum(abs(zt-forecast2)/((abs(zt)+abs(forecast2))/2))/length(zt)*100
write.table(forecast1,"D:/prediksiin7.txt")
write.table(forecast2,"D:/prediksiout7.txt")
rmse.in
rmse.out
smape.out

pre1=read.table("d:/predictas.txt",sep="," ,header=T)
pred1=predict(modpred,pre1)
pred1

```

Lampiran 21*Syntax SAS ARIMA (0,0,2)(0,1,1)³⁶ Pos Kedungadem*

```

data datakedung;
input x;
datalines;
0.072548
0.094072
0.125000
0.109109
0.156174
0.117851
0.112509
0.106000
:
:
0.258199
0.141421
0.102598
0.087039
0.125988
0.121268
0.258199
;
proc arima data=datakedung;
identify var=x(36);
estimate q=(1,2)(36) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 22*Syntax SAS ARIMA Outlier (0,0,2)(0,1,1)³⁶ Pos Kedungadem*

```

data datakedung;
input x;
datalines;
0.072548
0.094072
0.125000
:
:
0.087039
0.125988
0.121268
0.258199
;
data datakedung;
set datakedung;
if _n_=316 then ao1=1; else ao1=0;
if _n_=355 then ao2=1; else ao2=0;
if _n_=466 then ao3=1;else ao3=0;
if _n_=24 then ao4=1;else ao4=0;
if _n_=52 then ao5=1;else ao5=0;
run;
proc arima data=datakedung;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ao3(36) ao4(36)
ao5(36));
estimate q=(1,2)(36) input=(ao1 ao2 ao3 ao4 ao5) noconstant
method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 23**Output ARIMA Outlier (0,0,2)(0,1,1)³⁶ Pos Kedugadem**

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
MA1,1	-0.18551	0.04350	-4.26	<.0001	1	x	0
MA1,2	-0.14252	0.04333	-3.29	0.0011	2	x	0
MA2,1	0.73522	0.03039	24.19	<.0001	36	x	0
NUM1	-0.16004	0.04798	-3.34	0.0009	0	ao1	0
NUM2	0.15838	0.04811	3.29	0.0011	0	ao2	0
NUM3	0.14867	0.04840	3.07	0.0022	0	ao3	0
NUM4	-0.15915	0.03484	-4.57	<.0001	0	ao4	0
NUM5	-0.12753	0.04823	-2.64	0.0084	0	ao5	0

Autocorrelation Check of Residuals

To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----						
6	2.04	3	0.5646	0.005	0.012	0.039	0.037	-0.017	-0.019	
12	6.04	9	0.7359	0.027	-0.007	0.032	0.009	-0.067	0.029	
18	8.24	15	0.9137	0.028	0.032	0.012	-0.043	0.002	0.011	
24	15.03	21	0.8215	0.046	-0.052	0.045	0.020	-0.058	0.037	
30	16.71	27	0.9380	-0.020	-0.037	0.008	-0.022	0.007	-0.025	
36	19.36	33	0.9716	-0.018	-0.002	0.031	0.026	-0.051	-0.005	
42	22.95	39	0.9809	0.013	0.056	0.022	-0.045	-0.013	0.014	
48	29.79	45	0.9606	-0.020	-0.074	0.003	0.021	-0.045	-0.057	

Tests for Normality

Test	--Statistic--	-----p Value-----
Shapiro-Wilk	W 0.990833	Pr < W 0.0020
Kolmogorov-Smirnov	D 0.045094	Pr > D <0.0100
Cramer-von Mises	W-Sq 0.289556	Pr > W-Sq <0.0050
Anderson-Darling	A-Sq 1.6109	Pr > A-Sq <0.0050

Lampiran 24*Syntax* SAS ARIMA (1,0,1)(0,1,1)³⁶ Pos Kedungadem

```

data datakedung;
input x;
datalines;
0.072548
0.094072
0.125000
0.109109
0.156174
0.117851
0.112509
0.106000
:
:
0.258199
0.141421
0.102598
0.087039
0.125988
0.121268
0.258199
;
proc arima data=datakedung;
identify var=x(36);
estimate p=(1) q=(1)(36) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```


Lampiran 25*Syntax SAS ARIMA Outlier* (1,0,1)(0,1,1)³⁶ Pos Kedungadem

```

data datakedung;
input x;
datalines;
0.072548
0.094072
0.125000
0.109109
:
0.087039
0.125988
0.121268
0.258199
;
data datakedung;
set datakedung;
if _n_=466 then ao1=1; else ao1=0;
if _n_=316 then ao2=1; else ao2=0;
if _n_=355 then ao3=1;else ao3=0;
if _n_=265 then ao4=1;else ao4=0;
if _n_=359 then ao5=1;else ao5=0;
run;
proc arima data=datakedung;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ao3(36)
ao4(36) ao5(36));
estimate p=1 q=(1)(36) input=(ao1 ao2 ao3 ao4 ao5)
noconstant method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 26**Output ARIMA Outlier (1,0,1)(0,1,1)³⁶ Pos Kedungadem**

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
MA1,1	0.45681	0.14112	3.24	0.0013	1	x	0
MA2,1	0.70465	0.03168	22.24	<.0001	36	x	0
AR1,1	0.65191	0.12038	5.42	<.0001	1	x	0
NUM1	0.15441	0.04820	3.20	0.0014	0	ao1	0
NUM2	-0.15139	0.04780	-3.17	0.0016	0	ao2	0
NUM3	0.15890	0.04790	3.32	0.0010	0	ao3	0
NUM4	0.14841	0.04780	3.10	0.0020	0	ao4	0
NUM5	0.14280	0.04786	2.98	0.0030	0	ao5	0

Autocorrelation Check of Residuals

To Lag	Chi- Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	0.42	3	0.9359	-0.005	0.019	-0.014	-0.003	-0.007	-0.012
12	4.41	9	0.8827	0.014	0.011	0.053	-0.013	-0.063	-0.005
18	8.15	15	0.9175	0.053	-0.017	0.031	-0.043	-0.029	-0.004
24	12.57	21	0.9230	0.031	-0.036	0.047	0.045	-0.033	0.012
30	15.69	27	0.9586	-0.001	-0.069	0.017	0.020	0.007	0.006
36	19.71	33	0.9674	-0.001	0.004	0.025	0.014	-0.078	0.001
42	20.96	39	0.9919	0.003	0.041	0.014	0.002	-0.010	0.012
48	26.37	45	0.9879	-0.030	-0.045	0.020	-0.013	-0.034	-0.067

Tests for Normality

Test	--Statistic--		-----p Value-----	
Shapiro-Wilk	W	0.991374	Pr < W	0.0031
Kolmogorov-Smirnov	D	0.048105	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.324131	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1.726996	Pr > A-Sq	<0.0050

Lampiran 27*Syntax* SAS ARIMA (1,0,[1,52])(0,1,1)³⁶ Pos Kedungadem

```

data datakedung;
input x;
datalines;
0.072548
0.094072
0.125000
0.109109
0.156174
0.117851
0.112509
0.106000
:
:
0.258199
0.141421
0.102598
0.087039
0.125988
0.121268
0.258199
;
proc arima data=datakedung;
identify var=x(36);
estimate p=(1) q=(1,52)(36) noconstant method=cls;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 28*Syntax SAS ARIMA Outlier (1,0,[1,52])(0,1,1)³⁶ Pos Kedungadem*

```

data datakedung;
input x;
datalines;
0.072548
0.094072
0.125000
:
:
0.087039
0.125988
0.121268
0.258199
;
data datakedung;
set datakedung;
if _n_=466 then ao1=1; else ao1=0;
if _n_=316 then ao2=1; else ao2=0;
if _n_=355 then ao3=1;else ao3=0;
if _n_=265 then ao4=1;else ao4=0;
if _n_=359 then ao5=1;else ao5=0;
run;
proc arima data=datakedung;
identify var=x(36) crosscorr=(ao1(36) ao2(36) ao3(36) ao4(36)
ao5(36));
estimate p=1 q=(1)(36) input=(ao1 ao2 ao3 ao4 ao5) noconstant
method=cls;
outlier maxnum=5;
forecast out=ramalan lead=18;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc export data=ramalan
outfile="E:/Cawak.xls"
dbms=excel97
replace;
run;

```

Lampiran 29**Output ARIMA Outlier (1,0,[1,52])(0,1,1)³⁶ Pos Kedungadem**

Conditional Least Squares Estimation

Parameter	Estimate	Standard Error	t Value	Approx Pr > t	Lag	Variable	Shift
MA1,1	0.41924	0.13456	3.12	0.0019	1	x	0
MA1,2	-0.10144	0.04119	-2.46	0.0141	52	x	0
MA2,1	0.72463	0.03096	23.40	<.0001	36	x	0
AR1,1	0.62065	0.11685	5.31	<.0001	1	x	0
NUM1	0.16705	0.04767	3.50	0.0005	0	ao1	0
NUM2	0.15392	0.04738	3.25	0.0012	0	ao2	0
NUM3	0.14283	0.04720	3.03	0.0026	0	ao3	0
NUM4	-0.14276	0.04723	-3.02	0.0026	0	ao4	0
NUM5	-0.15653	0.03494	-4.48	<.0001	0	ao5	0

Autocorrelation Check of Residuals

To Lag	Chi-Square	DF	Pr > ChiSq	-----Autocorrelations-----					
6	1.42	2	0.4906	-0.012	0.035	-0.022	-0.000	-0.008	-0.026
12	5.94	8	0.6534	0.031	-0.003	0.037	0.034	-0.063	0.027
18	9.26	14	0.8140	0.032	0.004	0.020	-0.067	-0.004	0.002
24	16.31	20	0.6971	0.054	-0.040	0.051	0.030	-0.054	0.040
30	17.48	26	0.8939	-0.016	-0.030	0.016	-0.022	0.012	-0.005
36	21.61	32	0.9176	-0.016	0.004	0.023	0.026	-0.074	0.013
42	24.05	38	0.9619	0.010	0.053	0.014	-0.030	-0.013	-0.008
48	28.54	44	0.9657	-0.019	-0.056	0.005	0.006	-0.037	-0.051

Tests for Normality

Test	--Statistic--	-----p Value-----
Shapiro-Wilk	W 0.9891	Pr < W 0.0005
Kolmogorov-Smirnov	D 0.057297	Pr > D <0.0100
Cramer-von Mises	W-Sq 0.326565	Pr > W-Sq <0.0050
Anderson-Darling	A-Sq 1.903273	Pr > A-Sq <0.0050

Lampiran 30

Syntax Model SVR ARIMA $([1,16,37],0,0) (3,1,0)^{18}$ Pos
Kedungadem

```
library(e1071)
library(tseries)
data=read.csv("D:/kedung7.csv",sep=",",header=T)
attach(data)
train=data[1:467,]
test=data[-c(1:467),]
tuneResult=tune(svm, Yt
~,data=train,kernel="radial",ranges = list(epsilon =
seq(0.5,1), cost =seq(0.01,10),gamma=seq(1,1.5)))
tunedmodel=tuneResult$best.model;tunedmodel
forecast.in=predict(tunedmodel,train)
forecast.out=predict(tunedmodel,test)

forecast1=(forecast.in)
forecast2=(forecast.out)
testing=((test[,1]))
training=((train[,1]))
isi1=testing-forecast2
isi2=training-forecast1

rmse.in=sqrt(mean(isi2^2))
rmse.out=sqrt(mean(isi1^2))
smape.out=(sum(abs(testing-
forecast2)/((abs(testing)+abs(forecast2))/2))/length(testing))*100

write.table(forecast1,"D:/prediksisvrin.txt")
write.table(forecast2,"D:/prediksisvrout.txt")
rmse.in
rmse.out
smape.out
```

Lampiran 31

Hasil Penguraian Model ARIMA yang sudah memiliki parameter yang signifikan Pos Kedungadem.

ARIMA (2,0,0)

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_2 B^2) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* &= a_t \\ Y_t^* &= a_t + \phi_1 Y_{t-1}^* + \phi_2 Y_{t-2}^* \\ Y_t^* &= 0,5411 Y_{t-1}^* + 0,4264 Y_{t-2}^* + a_t \end{aligned}$$

ARIMA ([1,2,17],0,0)

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_2 B^2 - \phi_{17} B^{17}) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_{17} Y_{t-17}^* &= a_t \\ Y_t^* &= a_t + \phi_1 Y_{t-1}^* + \phi_2 Y_{t-2}^* + \phi_{17} Y_{t-17}^* \\ Y_t^* &= 0,5307 Y_{t-1}^* + 0,4049 Y_{t-2}^* + 0,0417 Y_{t-17}^* + a_t \end{aligned}$$

ARIMA ([8,10],0,0)

$$\begin{aligned} Y_t^* (1 - \phi_8 B^8 - \phi_{10} B^{10}) &= a_t \\ Y_t^* - \phi_8 Y_{t-8}^* - \phi_{10} Y_{t-10}^* &= a_t \\ Y_t^* &= a_t + \phi_8 Y_{t-8}^* + \phi_{10} Y_{t-10}^* \\ Y_t^* &= 0,5845 Y_{t-8}^* + 0,2889 Y_{t-10}^* + a_t \end{aligned}$$

ARIMA ([1,8],0,0)

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_8 B^8) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_8 Y_{t-8}^* &= a_t \\ Y_t^* &= a_t + \phi_1 Y_{t-1}^* + \phi_8 Y_{t-8}^* \\ Y_t^* &= 0,8123 Y_{t-1}^* + 0,1503 Y_{t-8}^* + a_t \end{aligned}$$

ARIMA (2,0,0)(2,1,0)¹⁸

$$\begin{aligned} Y_t^* (1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_{18} B^{18} - \Phi_{36} B^{36})(1 - B^{18}) &= a_t \\ Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \Phi_{18} Y_{t-18}^* + \phi_1 \Phi_{18} Y_{t-19}^* + \phi_2 \Phi_{18} Y_{t-20}^* \\ &\quad - \Phi_{36} Y_{t-36}^* + \phi_1 \Phi_{36} Y_{t-37}^* + \phi_2 \Phi_{36} Y_{t-38}^* - Y_{t-18}^* \\ &\quad + \phi_1 Y_{t-19}^* + \phi_2 Y_{t-20}^* + \Phi_{18} Y_{t-36}^* - \phi_1 \Phi_{18} Y_{t-37}^* \\ &\quad - \phi_2 \Phi_{18} Y_{t-38}^* + \Phi_{36} Y_{t-54}^* - \phi_1 \Phi_{36} Y_{t-55}^* \\ &\quad - \phi_2 \Phi_{36} Y_{t-56}^* = a_t \\ Y_t^* &= 0,3821 Y_{t-1}^* + 0,2547 Y_{t-2}^* - 0,7181 Y_{t-18}^* + 0,2746 Y_{t-19}^* \\ &\quad + 0,1831 Y_{t-20}^* - 0,1652 Y_{t-36}^* + 0,0631 Y_{t-37}^* \\ &\quad + 0,0421 Y_{t-38}^* + Y_{t-18}^* - 0,3821 Y_{t-19}^* \\ &\quad - 0,2547 Y_{t-20}^* + 0,7181 Y_{t-36}^* - 0,2746 Y_{t-37}^* \\ &\quad - 0,1831 Y_{t-38}^* + 0,1652 Y_{t-54}^* - 0,0631 Y_{t-55}^* \\ &\quad - 0,0421 Y_{t-56}^* + a_t \end{aligned}$$

ARIMA ([1,2,16],0,0)(3,1,0)¹⁸

$$Y_t^*(1 - \phi_1 B - \phi_2 B^2 - \phi_{16} B^{16})(1 - \Phi_{18} B^{18} - \Phi_{36} B^{36} - \Phi_{54} B^{54})(1 - B^{18}) = a_t$$

$$\begin{aligned} Y_t^* - \phi_1 Y_{t-1}^* - \phi_2 Y_{t-2}^* - \phi_{16} Y_{t-16}^* - \Phi_{18} Y_{t-18}^* + \phi_1 \Phi_{18} Y_{t-19}^* \\ + \phi_2 \Phi_{18} Y_{t-20}^* + \phi_{16} \Phi_{18} Y_{t-34}^* - \Phi_{36} Y_{t-36}^* \\ + \phi_1 \Phi_{36} Y_{t-37}^* + \phi_2 \Phi_{36} Y_{t-38}^* + \phi_{16} \Phi_{36} Y_{t-52}^* \\ - \Phi_{54} Y_{t-54}^* + \phi_1 \Phi_{54} Y_{t-55}^* + \phi_2 \Phi_{54} Y_{t-56}^* \\ + \phi_{16} \Phi_{54} Y_{t-70}^* - Y_{t-18}^* + \phi_1 Y_{t-19}^* + \phi_2 Y_{t-20}^* \\ + \phi_{16} Y_{t-34}^* + \Phi_{18} Y_{t-36}^* - \phi_1 \Phi_{18} Y_{t-37}^* \\ - \phi_2 \Phi_{18} Y_{t-38}^* - \phi_{16} \Phi_{18} Y_{t-52}^* + \Phi_{36} Y_{t-54}^* \\ - \phi_1 \Phi_{36} Y_{t-55}^* - \phi_2 \Phi_{36} Y_{t-56}^* - \phi_{16} \Phi_{36} Y_{t-70}^* \\ + \Phi_{54} Y_{t-72}^* - \phi_1 \Phi_{54} Y_{t-73}^* - \phi_2 \Phi_{54} Y_{t-74}^* \\ - \phi_{16} \Phi_{54} Y_{t-88}^* = a_t \end{aligned}$$

$$\begin{aligned} Y_t^* = 0,3301 Y_{t-1}^* + 0,2312 Y_{t-2}^* - 0,2826 Y_{t-16}^* - 0,6863 Y_{t-18}^* \\ + 0,2265 Y_{t-19}^* + 0,1587 Y_{t-20}^* - 0,1939 Y_{t-34}^* \\ - 0,4055 Y_{t-36}^* + 0,1339 Y_{t-37}^* + 0,0938 Y_{t-38}^* \\ - 0,1146 Y_{t-52}^* - 0,2460 Y_{t-54}^* + 0,0812 Y_{t-55}^* \\ + 0,0569 Y_{t-56}^* - 0,0695 Y_{t-70}^* - Y_{t-18}^* \\ - 0,3301 Y_{t-19}^* - 0,2312 Y_{t-20}^* + 0,2826 Y_{t-34}^* \\ + 0,6863 Y_{t-36}^* - 0,2265 Y_{t-37}^* - 0,1587 Y_{t-38}^* \\ + 0,1939 Y_{t-52}^* + 0,4055 Y_{t-54}^* - 0,1339 Y_{t-55}^* \\ - 0,0938 Y_{t-56}^* + 0,1146 Y_{t-70}^* + 0,2460 Y_{t-72}^* \\ - 0,0812 Y_{t-73}^* - 0,0569 Y_{t-74}^* + 0,0695 Y_{t-88}^* \\ + a_t \end{aligned}$$

ARIMA ([1,16,37],0,0)(3,1,0)¹⁸

$$Y_t^*(1 - \phi_1 B - \phi_{16} B^{16} - \phi_{37} B^{37})(1 - \Phi_{18} B^{18} - \Phi_{36} B^{36} - \Phi_{54} B^{54})(1 - B^{18}) = a_t$$

$$\begin{aligned} Y_t^* - \phi_1 Y_{t-1}^* - \phi_{16} Y_{t-16}^* - \phi_{37} Y_{t-37}^* - \Phi_{18} Y_{t-18}^* + \phi_1 \Phi_{18} Y_{t-19}^* \\ + \phi_{16} \Phi_{18} Y_{t-34}^* + \phi_{37} \Phi_{18} Y_{t-55}^* - \Phi_{36} Y_{t-36}^* \\ + \phi_1 \Phi_{36} Y_{t-37}^* + \phi_{16} \Phi_{36} Y_{t-52}^* + \phi_{37} \Phi_{36} Y_{t-73}^* \\ - \Phi_{54} Y_{t-54}^* + \phi_1 \Phi_{54} Y_{t-55}^* + \phi_{16} \Phi_{54} Y_{t-70}^* \\ + \phi_{37} \Phi_{54} Y_{t-91}^* - Y_{t-18}^* + \phi_1 Y_{t-19}^* + \phi_{16} Y_{t-34}^* \\ + \phi_{37} Y_{t-55}^* + \Phi_{18} Y_{t-36}^* - \phi_1 \Phi_{18} Y_{t-37}^* \\ - \phi_{16} \Phi_{18} Y_{t-52}^* - \phi_{37} \Phi_{18} Y_{t-73}^* + \Phi_{36} Y_{t-54}^* \\ - \phi_1 \Phi_{36} Y_{t-55}^* - \phi_{16} \Phi_{36} Y_{t-70}^* - \phi_{37} \Phi_{36} Y_{t-91}^* \\ + \Phi_{54} Y_{t-72}^* - \phi_1 \Phi_{54} Y_{t-73}^* - \phi_{16} \Phi_{54} Y_{t-88}^* \\ - \phi_{37} \Phi_{54} Y_{t-109}^* = a_t \end{aligned}$$

$$\begin{aligned}
Y_t^* = & 0,3865 Y_{t-1}^* - 0,2616 Y_{t-16}^* + 0,2108 Y_{t-37}^* - 0,7179 Y_{t-18}^* \\
& + 0,2775 Y_{t-19}^* - 0,1878 Y_{t-34}^* + 0,1513 Y_{t-55}^* \\
& - 0,4020 Y_{t-36}^* + 0,1554 Y_{t-37}^* - 0,1052 Y_{t-52}^* \\
& + 0,0847 Y_{t-73}^* - 0,2191 Y_{t-54}^* + 0,0847 Y_{t-55}^* \\
& - 0,0573 Y_{t-70}^* + 0,0462 Y_{t-91}^* - Y_{t-18}^* \\
& - 0,3865 Y_{t-19}^* + 0,2616 Y_{t-34}^* - 0,2108 Y_{t-55}^* \\
& + 0,7179 Y_{t-36}^* - 0,2775 Y_{t-37}^* + 0,1878 Y_{t-52}^* \\
& - 0,1513 Y_{t-73}^* + 0,4020 Y_{t-54}^* - 0,1554 Y_{t-55}^* \\
& + 0,1052 Y_{t-70}^* + 0,0847 Y_{t-91}^* + 0,2191 Y_{t-72}^* \\
& - 0,0847 Y_{t-73}^* + 0,0573 Y_{t-88}^* - 0,0462 Y_{t-109}^* \\
& + a_t
\end{aligned}$$

ARIMA (0,0,1)(0,1,1)³⁶

$$Y_t^* (1 - B^{36}) = a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36}$$

$$Y_t^* - Y_{t-36}^* = a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36}$$

$$Y_t^* = Y_{t-36}^* + a_t - \theta_1 a_{t-1} - \theta_{36} a_{t-36}$$

$$Y_t^* = Y_{t-36}^* + a_t + 0,1385 a_{t-1} - 0,7274 a_{t-36}$$

Lampiran 32

Syntax GA- SVR ARIMA Pos Kedungadem

```

library(e1071)
library(GA)
d=read.csv("d:/kedung7.csv",sep="," ,header=TRUE)
train_data = d[1:467,]
test_data = d[-c(1:467),]
# Fitness function (to be maximized)
# Parameter vector x is: (cost, gamma, epsilon)
fitnessFunc <- function(x) {
  # Retrieve the SVM parameters
  cost <- x[1]
  gamma <- x[2]
  epsilon <- x[3]
  # Train
  model <- svm(Yt ~ ., data = train_data, cost = cost, gamma = gamma,
epsilon = epsilon, type = "eps-regression", kernel = "radial")
  # Test
  rmse <- mean((predict(model, newdata = train_data) - train_data$Yt) ^ 2)
  return (-rmse)
}
# Range of the parameter values to be tested
# Parameters are: (cost, gamma, epsilon)
theta_min <- c(cost = 0.9, gamma = 0.9, epsilon = 0.4)
theta_max <- c(cost = 1.1, gamma = 1.1, epsilon = 0.6)
# Run the genetic algorithm
results <- ga(type = "real-valued", fitness = fitnessFunc,
  names = names(theta_min),
  min = theta_min, max = theta_max,
  popSize = 100, maxiter = 100, pcrossover=0.8, pmutation=0.01,
monitor=plot)
summary(results)
modpred=svm(Yt~.,data=train_data,cost=1.090917,gamma= 0.9028604
,epsilon=0.403149)
forecast.in=predict(modpred,train_data)
forecast.out=predict(modpred,test_data)
forecast1=(forecast.in)
forecast2=(forecast.out)
zt=(test_data[,1])
xt=(train_data[,1])
isi1=xt-forecast1
isi2=zt-forecast2
rmse.in=sqrt(mean(isi1^2))
rmse.out=sqrt(mean(isi2^2))
smape.out=sum(abs(zt-forecast2)/((abs(zt)+abs(forecast2))/2))/length(zt)*100
write.table(forecast1,"D:/prediksiin7.txt")
write.table(forecast2,"D:/prediksiout7.txt")
rmse.in
rmse.out
smape.out
pre1=read.table("d:/predictas.txt",sep="," ,header=T)
pred1=predict(modpred,pre1)
pred1

```

Lampiran 33

Data Curah Hujan Pos Cawak dan Kedungadem dari BMKG Karangploso Malang

**BADAN METEOROLOGI KLIMATOLOGI DAN GEOFISIKA
STASIUN KLIMATOLOGI KARANGPLOSO**



BNKG
Jl. Zentana No.33 Karangploso Malang
; Fax : (0341) 464827 ; Email : zentana33@yahoo.com , Website : karangploso.jatim.bnkg.go.id
; Fax : (0341) 464827, 481595
; Telp : (0341) 464827, 481595

DATA CURAH HUJAN TAHUN 2015 - 2016

Nama Stadion	: Cawak
No Stadion	: 3522090
Lintang	: 07° 13' 1
Bujur	: 112° 06'
Elevasi	: 20 m dpl

Desa : Simorejo
Kecamatan : Kepoh Baru
Kabupaten : Bojonegoro

Th	Jan	Feb	Mar	April	Mei	Jun	Jul	Agst	Sep	Okst	Nov	Des
2015	45	68	192	96	141	50	111	85	157	61	130	71
2016	48	40	115	72	218	14	46	112	71	66	65	10
2017	48	40	115	72	218	14	46	112	71	66	65	10
2018	48	40	115	72	218	14	46	112	71	66	65	10
2019	48	40	115	72	218	14	46	112	71	66	65	10
2020	48	40	115	72	218	14	46	112	71	66	65	10
2021	48	40	115	72	218	14	46	112	71	66	65	10
2022	48	40	115	72	218	14	46	112	71	66	65	10
2023	48	40	115	72	218	14	46	112	71	66	65	10
2024	48	40	115	72	218	14	46	112	71	66	65	10
2025	48	40	115	72	218	14	46	112	71	66	65	10
2026	48	40	115	72	218	14	46	112	71	66	65	10
2027	48	40	115	72	218	14	46	112	71	66	65	10
2028	48	40	115	72	218	14	46	112	71	66	65	10
2029	48	40	115	72	218	14	46	112	71	66	65	10
2030	48	40	115	72	218	14	46	112	71	66	65	10
2031	48	40	115	72	218	14	46	112	71	66	65	10
2032	48	40	115	72	218	14	46	112	71	66	65	10
2033	48	40	115	72	218	14	46	112	71	66	65	10
2034	48	40	115	72	218	14	46	112	71	66	65	10
2035	48	40	115	72	218	14	46	112	71	66	65	10
2036	48	40	115	72	218	14	46	112	71	66	65	10
2037	48	40	115	72	218	14	46	112	71	66	65	10
2038	48	40	115	72	218	14	46	112	71	66	65	10
2039	48	40	115	72	218	14	46	112	71	66	65	10
2040	48	40	115	72	218	14	46	112	71	66	65	10
2041	48	40	115	72	218	14	46	112	71	66	65	10
2042	48	40	115	72	218	14	46	112	71	66	65	10
2043	48	40	115	72	218	14	46	112	71	66	65	10
2044	48	40	115	72	218	14	46	112	71	66	65	10
2045	48	40	115	72	218	14	46	112	71	66	65	10
2046	48	40	115	72	218	14	46	112	71	66	65	10
2047	48	40	115	72	218	14	46	112	71	66	65	10
2048	48	40	115	72	218	14	46	112	71	66	65	10
2049	48	40	115	72	218	14	46	112	71	66	65	10
2050	48	40	115	72	218	14	46	112	71	66	65	10
2051	48	40	115	72	218	14	46	112	71	66	65	10
2052	48	40	115	72	218	14	46	112	71	66	65	10
2053	48	40	115	72	218	14	46	112	71	66	65	10
2054	48	40	115	72	218	14	46	112	71	66	65	10
2055	48	40	115	72	218	14	46	112	71	66	65	10
2056	48	40	115	72	218	14	46	112	71	66	65	10
2057	48	40	115	72	218	14	46	112	71	66	65	10
2058	48	40	115	72	218	14	46	112	71	66	65	10
2059	48	40	115	72	218	14	46	112	71	66	65	10
2060	48	40	115	72	218	14	46	112	71	66	65	10
2061	48	40	115	72	218	14	46	112	71	66	65	10
2062	48	40	115	72	218	14	46	112	71	66	65	10
2063	48	40	115	72	218	14	46	112	71	66	65	10
2064	48	40	115	72	218	14	46	112	71	66	65	10
2065	48	40	115	72	218	14	46	112	71	66	65	10
2066	48	40	115	72	218	14	46	112	71	66	65	10
2067	48	40	115	72	218	14	46	112	71	66	65	10
2068	48	40	115	72	218	14	46	112	71	66	65	10
2069	48	40	115	72	218	14	46	112	71	66	65	10
2070	48	40	115	72	218	14	46	112	71	66	65	10
2071	48	40	115	72	218	14	46	112	71	66	65	10
2072	48	40	115	72	218	14	46	112	71	66	65	10
2073	48	40	115	72	218	14	46	112	71	66	65	10
2074	48	40	115	72	218	14	46	112	71	66	65	10
2075	48	40	115	72	218	14	46	112	71	66	65	10
2076	48	40	115	72	218	14	46	112	71	66	65	10
2077	48	40	115	72	218	14	46	112	71	66	65	10
2078	48	40	115	72	218	14	46	112	71	66	65	10
2079	48	40	115	72	218	14	46	112	71	66	65	10
2080	48	40	115	72	218	14	46	112	71	66	65	10
2081	48	40	115	72	218	14	46	112	71	66	65	10
2082	48	40	115	72	218	14	46	112	71	66	65	10
2083	48	40	115	72	218	14	46	112	71	66	65	10
2084	48	40	115	72	218	14	46	112	71	66	65	10
2085	48	40	115	72	218	14	46	112	71	66	65	10
2086	48	40	115	72	218	14	46	112	71	66	65	10
2087	48	40	115	72	218	14	46	112	71	66	65	10
2088	48	40	115	72	218	14	46	112	71	66	65	10
2089	48	40	115	72	218	14	46	112	71	66	65	10
2090	48	40	115	72	218	14	46	112	71	66	65	10
2091	48	40	115	72	218	14	46	112	71	66	65	10
2092	48	40	115	72	218	14	46	112	71	66	65	10
2093	48	40	115	72	218	14	46	112	71	66	65	10
2094	48	40	115	72	218	14	46	112	71	66	65	10
2095	48	40	115	72	218	14	46	112	71	66	65	10
2096	48	40	115	72	218	14	46	112	71	66	65	10
2097	48	40	115	72	218	14	46	112	71	66	65	10
2098	48	40	115	72	218	14	46	112	71	66	65	10
2099	48	40	115	72	218	14	46	112	71	66	65	10
2100	48	40	115	72	218	14	46	112	71	66	65	10

Nama Stasiun : Panjang/ Kedung Adem
No Stasiun : 35220801a
Lintang : 07° 19' 51.4" LS
Bujur : 112° 03' 00.8" BT
Elevasi : 78 m dpl

Desa : Panjang
Kecamatan : Kedungadem
Kabupaten : Bojonegoro

Th	Jan	Feb	Mar	April	May	June	July	Aug	Sep	Oct	Nov	Dec
1	2	3	1	2	3	1	2	3	1	2	3	1
2	3	1	2	3	4	2	3	4	2	3	4	2
3	4	2	3	4	5	3	4	5	3	4	5	3
4	5	3	4	5	6	4	5	6	4	5	6	4
5	6	4	5	6	7	5	6	7	5	6	7	5
6	7	5	6	7	8	6	7	8	6	7	8	6
7	8	6	7	8	9	7	8	9	7	8	9	7
8	9	7	8	9	10	8	9	10	8	9	10	8
9	10	8	9	10	11	9	10	11	9	10	11	9
10	11	9	10	11	12	10	11	12	10	11	12	10
11	12	10	11	12	13	11	12	13	11	12	13	11
12	13	11	12	13	14	12	13	14	12	13	14	12
13	14	12	13	14	15	13	14	15	13	14	15	13
14	15	13	14	15	16	14	15	16	14	15	16	14
15	16	14	15	16	17	15	16	17	15	16	17	15
16	17	15	16	17	18	16	17	18	16	17	18	16
17	18	16	17	18	19	17	18	19	17	18	19	17
18	19	17	18	19	20	18	19	20	18	19	20	18
19	20	18	19	20	21	19	20	21	19	20	21	19
20	21	19	20	21	22	20	21	22	20	21	22	20
21	22	20	21	22	23	21	22	23	21	22	23	21
22	23	21	22	23	24	22	23	24	22	23	24	22
23	24	22	23	24	25	23	24	25	23	24	25	23
24	25	23	24	25	26	24	25	26	24	25	26	24
25	26	24	25	26	27	25	26	27	25	26	27	25
26	27	25	26	27	28	26	27	28	26	27	28	26
27	28	26	27	28	29	27	28	29	27	28	29	27
28	29	27	28	29	30	28	29	30	28	29	30	28
29	30	28	29	30	31	29	30	31	29	30	31	29
30	31	29	30	31		30	31		30	31		30
31												

ing, 29 Desember 2016

Unit pelayanan
Stasiun Kumpulologi Karangploso Malang



NO WUJIANBARI
97901112000122001

BIODATA PENULIS



Penulis yang bernama lengkap Kiki Wulan Suci adalah putri ketiga dari pasangan Khayin Abdad dan Siti Aisyah, lahir di Gresik pada tanggal 2 Februari 1995. Penulis menempuh jenjang pendidikan TK di TK Nurul Huda Leran, kemudian jenjang SD di MI Nurul Huda Leran. Pada jenjang SMP, penulis bersekolah di SMPN 1 Bungah dan melanjutkan jenjang SMA di SMAN 1

Manyar. Pada tahun 2013, penulis diterima sebagai mahasiswa S1 jurusan Statistika ITS melalui jalur SBMPTN. Selama menempuh jenjang S1, penulis bergabung sebagai anggota staff AD dan Tim Ahli Divisi PST HIMASTA-ITS. Selama 1 tahun penulis juga bergabung menjadi anggota di UKM ITS *Badminton Community* (IBC). Penulis juga pernah mengikut seminar *Inter-national Conference On Theoretical and Applied Statistics* (IC-TAS) yang diselenggarakan oleh Statistika-ITS tahun 2016. Apa-bila pembaca ingin berdiskusi mengenai tugas akhir ini dapat menghubungi penulis melalui alamat email yakni kwulansuci@gmail.com atau nomor telepon (085608761360).

(Halaman ini sengaja dikosongkan)